



Department of Health and Human Services
Centers for Disease Control and Prevention



**Cancer Surveillance Branch
National Program of Cancer Registries—
Advancing E-cancer Reporting and Registry
Operations (NPCR-AERRO)‡**

**NPCR-AERRO: Developing a Cancer
Surveillance Informatics Structure in
the New E-Health Environment**

April 2009

Contents

1	Executive Summary	4
2	Background.....	5
2.1	The DCPC/NPCR Mandate	5
2.2	Cancer Surveillance Infrastructure	5
2.3	National Health-IT Infrastructure	9
3	Problem Statement	10
3.1	Delay in Availability of Data	10
3.2	Resources for Collecting Data	10
3.3	Completeness of Reporting	11
3.4	Standardized Data Exchange for Non-Cancer Registry Data Sources	11
3.5	Limited data set	12
4	NPCR-AERRO Project Overview	13
4.1	Purpose.....	13
4.2	Scope of NPCR-AERRO	13
4.3	Context of Clinical Care Data and Cancer Surveillance	16
4.4	NPCR-AERRO Stakeholders.....	17
4.5	Constraints.....	18
4.6	Project Approach	21
5	Modeling	22
5.1	Modeling Activities	23
5.2	Technical Approach for Modeling	30
6	Analysis and Design	34
6.1	Analysis and Design Activities	34
7	Implementation	37
7.1	Implementation Activities	38
8	Project Management	41
8.1	Methodology	41
8.2	Development Phases.....	41
8.3	Timeline	43
9	Resources	44
10	Appendices	45
	Appendix 1: Domain Diagrams	45
	Appendix 2: Use Case Diagrams	48
11	References	53

Figures

Figure 1: NPCR-AERRO Context Diagram.....	14
Figure 2: NPCR-AERRO Scope Diagram.....	15
Figure 3: NPCR-AERRO Approach	21
Figure 4: NPCR-AERRO Process Flow for Modeling	22
Figure 5: Hospital and Central Cancer Registry Activity Plan	24
Figure 6: Cancer Control and Data Use Activity Plan	28
Figure 7: National and International Activity Plan	29
Figure 8: NPCR-AERRO Process Flow for Analysis and Design	34
Figure 9: NPCR-AERRO Process Flow for Implementation	37
Figure 10: Electronic Pathology Reporting Activity Plan	38
Figure 11: NPCR-AERRO Hospital: Domain Diagram.....	46
Figure 12: NPCR-AERRO CCR: Domain Diagram.....	47
Figure 13: NPCR-AERRO Hospital: Business Use Case Diagram.....	49
Figure 14: NPCR-AERRO Hospital: Operations Use Case Diagram.....	50
Figure 15: NPCR-AERRO CCR: Business Use Case Diagram.....	51
Figure 16: NPCR-AERRO CCR: Operations Use Case Diagram.....	52

Tables

Table 1: Data Collection Process.....	10
Table 2: EMR, EHR and PHR Definitions	16
Table 3: NPCR-AERRO Stakeholders.....	17
Table 4: Hospital Cancer Registry Core Functions.....	24
Table 5: Potential Impact of NPCR-AERRO Model	25
Table 6: Central Cancer Registry Core Functions	26
Table 7: NPCR-AERRO Diagrams	33

1 Executive Summary

The National Program of Cancer Registries (NPCR) is funded and managed by the Cancer Surveillance Branch (CSB), Division of Cancer Prevention and Control (DCPC) of the Centers for Disease Control and Prevention (CDC). CDC builds state and national capacity through support of the NPCR to monitor the burden of cancer including disparities among various subgroups in the population and provides data for research; evaluation of cancer control activities; and planning for future healthcare needs.

Cancer surveillance presents several challenges, including delay in availability of data, resources for collecting data, completeness of reporting, lack of standardized data exchange for non-cancer registry data sources, and limited datasets.

To help in addressing these needs, the CSB supports the NPCR Advancing E-cancer Reporting and Registry Operations (NPCR-AERRO) project which develops best practices, guidelines, and recommendations for an ideal cancer surveillance informatics infrastructure that takes advantage of emerging health information technology and national and international standards. NPCR-AERRO uses a collaborative framework to construct a comprehensive model to demonstrate the potential of electronic cancer registry reporting and automated registration to grantees and partners.

Goals in addressing challenges with cancer surveillance data collection include:

- Improve completeness, timeliness, and quality of data.
- Reduce costs for registries and data providers significantly over time.
- Develop a national plan or “blueprint” to identify priorities that make better use of cancer surveillance resources.
- Provide guidance for development of standards based systems for cancer registries.
- Improve data exchange between systems through use of industry standards.

NPCR-AERRO works with hospitals, central cancer registries, and national programs and data sources such as pathology laboratories, hospital registries, and physician offices to help in meeting these goals in the context of the emerging development and use of the Electronic Medical Record (EMR), Electronic Health Record (EHR), and Personal Health Record (PHR).

NPCR-AERRO focuses on a three-pronged approach:

- **Modeling:** Consensus best practices for electronic cancer registration and reporting
- **Analysis/Design:** Requirements for data and tools to further define best practices
- **Implementation:** Demonstrated use of best practice models and requirements in cancer registry software, training tools, and related products

NPCR-AERRO relies on standard information technology tools and processes including Unified Modeling Language (UML) and the CDC Unified Process (CDC UP) to develop models, requirements, and implementation products, and to manage the project as a whole.

Overall, NPCR-AERRO demonstrates how the EHR, consensus standards, electronic reporting and automation can improve the timeliness, completeness and quality of cancer registry data.

All NPCR-AERRO activities are scoped and tracked by approach. The Cancer Surveillance Branch website at <http://www.cdc.gov/cancer/npcr/informatics/merp/index.htm> provides up-to-date descriptions and status of project activities.

The website also houses a “CyberView” library of best practice models:
<http://www.cdc.gov/cancer/npcr/informatics/merp2/>

2 Background

2.1 The DCPC/NPCR Mandate

In direct response to the need for monitoring the cancer burdens for over one million Americans that are diagnosed with a reportable neoplasm annually, the CDC was authorized to establish the National Program of Cancer Registries (NPCR) program.¹

The Cancer Surveillance Branch within the Division of Cancer Prevention and Control (DCPC), Centers for Disease Control and Prevention (CDC) provides funds and technical assistance to improve cancer registration and cancer surveillance throughout the United States. CDC's goal is to build state and national capacity through the NPCR in order to monitor the burden of cancer including the disparities among various subgroups in the population; and provide data for research and evaluation of cancer control activities; and plan for future healthcare needs.

In October 1992, in order to address the need for cancer incidence data for planning and evaluating cancer control activities, Congress passed Public Law 102-515², the Cancer Registries Amendment Act. The Cancer Registries Amendment Act became a major milestone for cancer registration in the United States. This law established the NPCR and authorized CDC to:

- Provide funds and technical assistance to improve or enhance existing statewide central cancer registries, or plan and implement statewide central cancer registries where they did not already exist.
- Set standards for data completeness, timeliness, and quality.
- Assist with development of model laws and regulations for states and territories authorizing cancer registries and enhancing their viability of operations.
- Establish a set of required data items and a uniform standard reporting format.
- Provide training and support related to central registry operations.

Public Law 107-260, the Benign Brain Tumor Cancer Registries Amendment Act³, requires programs participating in the NPCR to collect data on benign and borderline tumors of the central nervous system in addition to the previously required data on malignant tumors. In 1998 Congress passed additional legislation to reauthorize the program.

2.2 Cancer Surveillance Infrastructure

Disease surveillance is undergoing dramatic changes with the potential to play a key role in improved healthcare delivery and public health practices. Cancer surveillance is a valuable tool in the arena of chronic disease management and cancer control. Within the United States the issues of completeness, timeliness, and quality of data prove to be a continual challenge. These issues become more evident as the agenda for a National Health Information Network (NHIN)⁴ gains momentum. Such national efforts eventually call for innovation at all levels of the healthcare delivery and public health infrastructure, and require high degrees of coordination and integration of resources.

The cancer surveillance infrastructure consists of a complex network of hospitals, clinics, laboratories, health departments, non-governmental organizations, and government agencies. This network is infused with professionals of varying levels of responsibility and accountability for accurate measures of cancer incidence throughout the country.

In addition to recording the occurrence of each case of cancer, the reporters provide information on the diagnosis, treatment and outcomes. Monitoring, recording and consolidating information on people diagnosed with cancer within the national cancer surveillance infrastructure ultimately leads to accurate and complete data on cancer incidence.

Data from population-based cancer registries can be used to:

- Monitor cancer trends over time.
- Determine cancer patterns in various populations.
- Provide complete state and national cancer incidence (i.e. number of cancers diagnosed in a population).
- Provide cancer information to the healthcare community and the public;
- Guide comprehensive cancer control planning and evaluation of cancer control programs (e.g., determine whether prevention, screening, and treatment efforts are making a difference).
- Help set priorities for allocating health resources.
- Advance clinical, epidemiologic, and health services research.⁵

Data from hospital-based cancer registries provides information locally that leads to the improvement of patient care within the clinical care community. Hospital registry data can be used in evaluation and quality initiatives, and can stimulate improvement in diagnosis and treatment. Hospitals contribute collectively to the population-based or state/territory-based registries (central cancer registries) which collect data for a specified geographic region.

2.2.1 Hospital Community

The lowest level of aggregation occurs within the hospital as the patient interacts with the healthcare system. Such interactions involve patient entry (in-patient/out-patient admission), diagnostic testing, treatment, counseling, etc. Several key departments within the healthcare system are involved in the process and serve as key points of interest within the NPCR-AERRO effort. These include but are not limited to the laboratory services, claims department, and healthcare providers, all of which interact with the patient at some point throughout the continuum of cancer care. Such in-hospital departments (sources) typically prepare reports that are submitted to/extracted by the hospital cancer registry, making it a key data aggregation point.

Hospital cancer registries send aggregated data reports or cancer abstracts to a central cancer registry at the state/regional level. Hospitals that have a Cancer Center approved by the American College of Surgeons (ACoS) Commission on Cancer's (CoC) also send their de-identified cases to the National Cancer Data Base (CoC/NCDB). Given the fact that a patient's care may be given in multiple institutions, there can be data sharing that takes place between hospital cancer registries. Additionally, state laws require hospitals without a cancer registry to report cancers (and other reportable tumors) to the state or regional registry.⁶ The mechanism used by hospitals without cancer registries to report to the state or regional registry varies depending on the size of the hospital, staff resources, and the reporting laws and/or policies of the respective states.

2.2.2 State/Regional Community

Similar to the hospital level, there is a need for data aggregation at the state or regional level. While at the hospital level data aggregation is dedicated to the patient population within its system(s) of care, the state/regional level aggregation spans all healthcare facilities within a

given population (i.e., state or region). As such, the scale of aggregation is increased exponentially from the hospital to the central cancer registry to provide population-based data. Central registry data sources can be thought of in two distinct categories, hospital and non-hospital sources of data. At the state/regional level, the hospital submits data using the hospital-based cancer abstract. In contrast, data sources like the free-standing or reference pathology laboratory and the healthcare provider/clinician are common types of non-hospital sources of cancer data. These and other non-hospital sources also report data to a central cancer registry. Pathology laboratories submit the actual cancer pathology report. The central cancer registry may also share information with the hospital cancer registry in accordance with state laws and, in addition, reports de-identified data to organizations at the national level.

2.2.3 National Community

Federally funded central cancer registries are required to report aggregate data to one or more national programs. These national population-based cancer programs represent the final level of reporting and aggregation of cancer occurrence in the United States to produce the annual report to the nation, the United States Cancer Statistics (USCS)⁷ publication and data for cancer control (e.g., Cancer Control Planet⁸).

2.2.3.1 Centers for Disease Control and Prevention's (CDC's) National Program of Cancer Registries (NPCR)

Established by Congress through the Cancer Registries Amendment Act in 1992, and administered by the CDC, NPCR collects data on the occurrence of cancer; the type, extent, and location of the cancer; and the type of initial treatment. These data represent 98 percent of the U.S. population. (See Section 2.1 for more information on the NPCR). The website for the NPCR can be found at: www.cdc.gov/cancer/npcr/.

2.2.3.2 National Cancer Institute's (NCI's) Surveillance, Epidemiology and End Results (SEER) Program

The SEER Program of the NCI is an important source of information on cancer incidence and survival in the United States. SEER currently collects and publishes long-term cancer incidence and survival data from population-based cancer registries covering up to approximately 26 percent of the U.S. population. The SEER Program produces an annual *Cancer Statistics Review* (CSR)⁹ which shows cancer statistics by race, sex, age, and year of diagnosis for the major cancer sites and for all cancers combined. The website for the SEER Program can be found at: www.seer.cancer.gov.

In addition to these funding agencies, cancer registries may submit data to standard setting organizations

2.2.3.3 North American Association of Central Cancer Registries (NAACCR)

NAACCR is a collaborative umbrella organization for cancer registries, governmental agencies, professional organizations, and private groups in North America interested in enhancing the quality and use of cancer registry data. NAACCR maintains a central registry certification program that annually reviews member registries for their ability to produce complete, accurate, and timely data. NAACCR produces *Cancer in North America* (CINA)¹⁰ annually to provide cancer incidence and mortality statistics for the United States and Canada. The website for NAACCR can be found at: www.naacr.org.

2.2.3.4 American College of Surgeons (ACoS) Commission on Cancer's (CoC) National Cancer Data Base (NCDB)

The CoC is a consortium of professional organizations dedicated to improving survival and quality of life for cancer patients through standard-setting, prevention, research, education, and the monitoring of comprehensive quality care. It performs analysis of Hospital Registry Data. Data from accredited hospital-based cancer registries are reported to the CoC/NCDB. Although this database is not population-based, there are more treatment data items available than in the population-based cancer registry programs. The website for the Commission on Cancer can be found at: www.facs.org/cancer/index.html.

2.2.4 International Community

Cancer registration is an internationally recognized requirement for effective cancer control. The World Cancer Declaration, issued in July 2006 by the International Union Against Cancer (UICC) at the World Cancer Congress, contained 11 specific actions for the global cancer control community, one of which was to "Increase the number of countries with viable and adequately funded cancer surveillance systems, including cancer registries."¹¹

2.2.4.1 International Association for Research on Cancer (IARC)

The World Health Organizations International Association for Research on Cancer provides, coordinates, and conducts research on cancer, and works to develop methods for cancer control. The IARC is also responsible for producing the internationally recognized reference source Cancer in Five Continents (CI5), documenting the incidence of cancer in populations around the world. Published every five years, it describes the burden of cancer based on original data collected by population-based cancer registries. Data from NPCR, as well as several of its participating regional and stage cancer registries, are included in CI5.¹² The website for the IARC can be found at: www.iarc.fr/.

2.2.4.2 International Association of Cancer Registries (IACR)

The International Association of Cancer Registries (IACR) is a professional society primarily for population registries to foster the aims and activities cancer registries internationally to improve the quality of data and comparability between registries. The NPCR-AERRO is collaborating with IACR to develop methods for automated cancer registration in order to meet the increasing need for more detailed, consistent and timely cancer data.¹³ The website for the IACR can be found at: www.iacr.com.fr/.

2.2.4.3 International Health Terminology Standards Development Organisation (IHTSDO)

Systematized Nomenclature of Medicine--Clinical Terms (SNOMED CT®) is a comprehensive clinical terminology, originally created by the College of American Pathologists (CAP) and, as of April 2007, owned, maintained, and distributed by the IHTSDO, a not-for-profit association in Denmark.

SNOMED CT® is one of a suite of designated standards for use in U.S. Federal Government systems for the electronic exchange of clinical health information and is also a required standard in interoperability specifications of the U.S. Healthcare Information Technology Standards Panel (HITSP). SNOMED CT® is also being implemented internationally as a standard within other IHTSDO Member countries.¹⁴ The website for the IHTSDO can be found at: www.ihtsdo.com.

2.3 National Health-IT Infrastructure

During the State of the Union Address on January 20, 2004, President George W. Bush stated that “By computerizing health records, we can avoid dangerous medical mistakes, reduce costs, and improve care.”¹⁵ The President set a health initiative goal to implement an Electronic Health Record (EHR) within 10 years that would be available for most Americans at any time or place. In order to achieve this goal, the President charged the Secretary of Health with overseeing activities that would transition the United States healthcare system from a traditional paper-based system to standardized electronic health records. The Secretary of Health created the Office of the National Coordinator (ONC) at the U.S. Department of Health and Human Services (HHS) to serve as his advisor and provide leadership for the development and implementation of a nationwide health information network.

The ONC facilitates several activities that are moving the nation towards a nationwide EHR. Activities include:

- Forming a federally-chartered advisory committee, the American Health Information Community (AHIC), that makes recommendations to the Secretary of HHS on how to create health records that are digital and interoperable, encourage market-led adoption, and ensure that the privacy and security of those records are protected at all times.
- Forming HITSP, a partnership between public and private sectors with the common goal of developing a widely accepted and useful set of standards enabling interoperability among healthcare software applications at the local, regional and national health information network in the United States. HITSP responds to requests from HHS and AHIC.
- NHIN is an initiative to develop a nationwide interoperable health information infrastructure that will allow the secure exchange of information across the healthcare community, including the consumer. This network will enable consumers and healthcare providers to have ready access to their health information at any location or point in time for clinical decision making.¹⁶

The Healthcare Information and Management Systems Society (HIMSS)-sponsored Integrating the Healthcare Enterprise (IHE) fosters participation and collaboration among healthcare professionals and industry to find ways to improve data exchange between healthcare computer systems. IHE supports the use of national established standards such as Digital Imaging and Communication in Medicine (DICOM) and Health Level Seven (HL7). There is cross-fertilization between the HITSP and IHE initiatives. These two initiatives actually utilize and build upon work from each other.

The Public Health Data Standards Consortium (PHDSC) was established in 1999 to promote the use of standardized information on health and healthcare. The PHDSC was incorporated in 2003 as a not-for-profit organization. It is a national non-profit member-based partnership of federal, state and local health agencies; national and local professional associations; and public and private sector organizations and individuals. The PHDSC is committed to bringing a common voice from public health and health services research communities to the national data standardization efforts.

CDC through NPCR-AERRO provides a service to the cancer community by participating in these national health-IT initiatives to ensure that the needs of public health and cancer surveillance are addressed. NPCR-AERRO also applies standards defined and/or adopted by these national health-IT initiatives in all project activities.

3 Problem Statement

Cancer surveillance is a very complex system that captures longitudinal data from multiple and varying data sources using a variety of methods. Data collection standards for reporting cancer data from hospital cancer registries to central cancer registries and then to the national cancer programs have existed for many years. NPCR-AERRO addresses problems that are outlined below in a variety of ways. The methods for addressing these issues are described in Section 4.6 with specific activity detailed provided in Sections 5, 6, and 7.

3.1 Delay in Availability of Data

The time gap between a diagnosis of cancer and the availability of data for analysis is a significant problem for cancer surveillance. Data from state cancer registries are generally transmitted to national programs many months after the diagnosis year.¹⁷ As a result, the annual national publications reflect information about patients diagnosed more than two years earlier. For example, for patients diagnosed in 2006, data are unavailable until April, 2009. Cancer prevention and control programs and others using the data would like data on a more rapid basis in order to more effectively evaluate and react to trends.¹⁸

The lag time between diagnosis of the cancer and reporting from a hospital to the central registry is generally 7 to 12 months, following the standard set by the American College of Surgeons' Commission on Cancer for providing cancer information.¹⁹ While incremental advancements in electronic methodology are being implemented, most hospital registries rely on traditional manual methods to identify reportable tumors and to abstract salient information into electronic data collection systems.

3.2 Resources for Collecting Data

The process of identifying and collecting cancer data is resource intensive, time consuming, and creates a risk of errors in transcription. The California Cancer Registry reviewed 12,116 pathology reports from 2006 to evaluate the time and cost savings between manual and electronic/automated review. 5,200 routine cases and 2,908 consultation cases were identified.²⁰ *Table 1: Data Collection Process* shows the benefit of electronic review.

Table 1: Data Collection Process

Process	Manual	E-Path	Savings
Number of hours to review and sort	808	707	\$10,100
Number of hours to enter 5,200 cases into registry database	260	0	\$6,500
Number of hours to enter 2,908 Consult cases into registry database	145	0	\$3,625
Copying pathology reports \$.10/report + Paper (\$2.55/ream)	Yes	No	\$1,274
TOTAL Savings			\$21,499

The logistical burden of reporting and the restriction of receiving electronic updates by the central registry have meant that hospital cancer registries submit a single report, and only when

the full first course of treatment has been started. For many cases, the time needed to collect data on all first course of treatment is quite lengthy.²¹

Central cancer registries have the burden of linking multiple records for a patient that relate to a single cancer, validating the consolidated data, linking to other data sources such as geo-coding databases and the Indian Health Service to improve the data, and performing death clearance activities to obtain vital status for survival analyses. The timeline for performing these activities can be impacted by external sources. For example, the availability of annual mortality data can be delayed as much as 6-12 months. National cancer programs set their submission timeline requirements taking the delay into account.

3.3 Completeness of Reporting

Complete and high quality cancer reporting has until recent years been achievable by primarily relying on hospital cancer registries. Traditionally cancer patients receive diagnostic testing or work-up and/or treatment in hospitals. Advances in medicine now allow patients to obtain their care outside the acute care hospital setting. Data collection systems from other sources such as physician offices and radiation therapy centers, however, are not as consistent with reporting. This leads to under reporting of certain types of cancers, typically those now diagnosed and treated outside the acute care hospital setting. Both melanomas and prostate cancers, for example have been shown to be under reported when central registries rely only on hospital reporting.^{22, 23}

In many states, non-hospital data sources are only minimally involved in reporting to the central cancer registry although the numbers are increasing each year. When reporting does occur, it may be through a manual process of identifying reportable cases and submitting copies of the medical record or by the central registry sending certified tumor registrars (CTR) into the facilities and physicians offices to manually abstract the information from the paper-based medical records. These processes are very resource intensive, time-consuming, and vulnerable to errors in transcription.

3.4 Standardized Data Exchange for Non-Cancer Registry Data Sources

Many of the state central cancer registries have worked independently to develop methods to receive critical data from hospitals without a cancer registry and non-hospital health care reporting sources, resulting in a variety of data collection, transmission, and reporting systems (or tools). The lack of coordination across the cancer and healthcare communities for seamless data exchange has contributed to incomplete reporting of cancers from non-hospital sources.

The emphasis on increased data timeliness and continued completeness highlights the growing need for standardization and automation for collecting and reporting critical cancer data from across the healthcare community. This emphasis has the potential for positive impact on the already high quality data. The cancer registry community has not yet taken full advantage of the growing number of healthcare facilities adopting an electronic health record (EHR) system to collect and report cancer cases more rapidly and completely.²⁴ Additionally, national efforts on electronic data exchange and processing are occurring which impact cancer registry operations.²⁵

Registries have used multiple data sets to improve the quality of their data. Vital statistics, health claims data, voter registration and many others contribute specific pieces of data that complete the cancer registry database. Standard data exchange for these data sets is lacking,

requiring each central registry to develop specific record layout formats for each data set in their state.

3.5 Limited data set

Cancer surveillance has had to limit its data set due to the expense of manually collecting data. More specific information, such as the drug(s) given for hormonal or chemotherapy, or socio-economic factors such as income or education level, could be of use for public health planners, analysts and epidemiologists, but is too labor intensive to collect manually. In some cases, these variables have been collected in special studies.

4 NPCR-AERRO Project Overview

4.1 Purpose

The purpose of the National Program of Cancer Registries – Advancing E-cancer Reporting and Registry Operations (NPCR-AERRO) is to develop best practices, guidelines, and recommendations for a cancer surveillance infrastructure that takes advantage of emerging health information technology and national and international standards. Emerging technologies include the Electronic Health Record (EHR), eHealth initiatives, and cancer registration informatics activities. The project focuses on three distinct organizational environments: 1) the hospital environment; 2) a state-wide, population-based central cancer registry; and 3) a national cancer program.

The purpose of NPCR-AERRO is to construct a comprehensive model that can enable CDC to demonstrate the potential of electronic cancer registry reporting and automated registration to its grantees and partners. NPCR-AERRO is a project funded by the Cancer Surveillance Branch within CDC's Division of Cancer Prevention and Control that involves a collaboration of public and private sector organizations committed to the idea of automating cancer registry operations for the purpose of increasing completeness, timeliness, and quality of data used to accurately articulate the national cancer burden.

The Cancer Surveillance Branch's strategic goals are to:

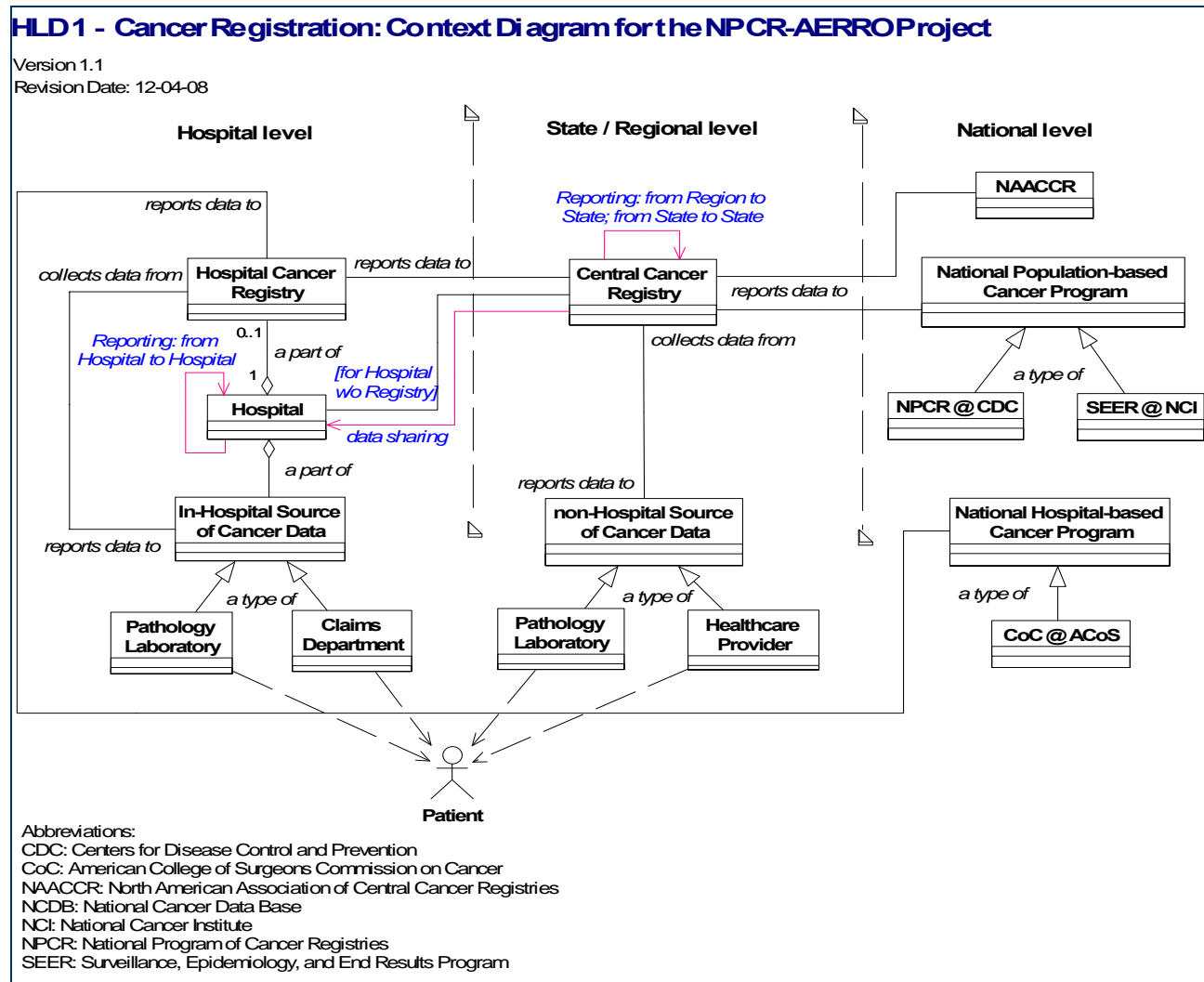
- Improve completeness, timeliness, and quality of data.
- Reduce costs for registries and data providers significantly over time.
- Develop a national plan or "blueprint" to identify priorities that make better use of cancer surveillance resources.
- Provide guidance for development of standards based systems for cancer registry.
- Improve data exchange between systems through use of industry standards.

To assist in meeting these goals NPCR-AERRO works to:

- Identify new capabilities offered by electronic capture of patient information.
- Identify opportunities to automate manual processes for data capture.
- Incorporate national standards.
- Reflect current industry best practices.
- Use iterative process to develop and assess models, design specifications, and implementation at multiple levels of granularity and specificity.

4.2 Scope of NPCR-AERRO

NPCR-AERRO concentrates on reporting and registry operations at three levels: the hospital, state/regional, and national. *Figure 1: NPCR-AERRO Context Diagram* shows the interactions between the patient and the different healthcare settings. NPCR-AERRO focuses on ways to enhance data reporting from multiple data sources at the different cancer registry levels and identifies areas that can benefit from the implementation of automated cancer registry operations.

Figure 1: NPCR-AERRO Context Diagram

NPCR-AERRO focuses on all current and potential data sources for the hospital and central cancer registry levels. Hospitals are the first priority because they provide the majority of cancer data. Electronic reporting of hospital data can improve the timeliness, completeness, and quality of cancer surveillance data reported at the state and national levels. Other data sources, such as pathology and physicians' offices, are important to address as well.

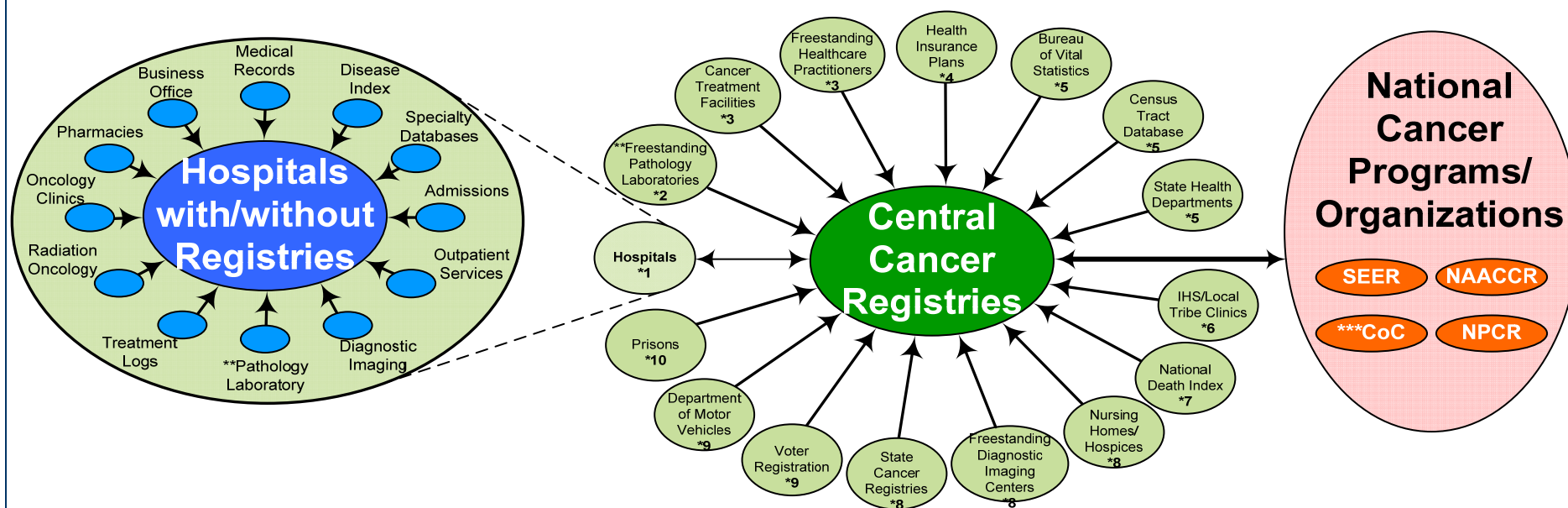
Figure 2: NPCR-AERRO Scope Diagram is a simple flow diagram that identifies the multiple data sources in a ranked order, based on registries' experience of the quantity of useful data that are available and reported to the central cancer registry. This diagram provides a simplified high-level view of the project scope for the hospital and central cancer registry levels.

Figure 2: NPCR-AERRO Scope Diagram

NPCR – AERRO Scope Diagram

Version 1.0

Revised 12-11-2008



NPCR-AERRO includes cancer data sources and the lines drawn to the Central Cancer Registries and the National Cancer Programs

Numbers rank the data sources on the quality of useful data available on a scale of 1 being the most useful and 10 being the least useful.

*Pathology Laboratories—Freestanding and Hospital—send data to both the Hospital Registries and the Central Cancer Registries

**CoC receives data directly from hospitals.

4.3 Context of Clinical Care Data and Cancer Surveillance

NPCR-AERRO and all other cancer informatics efforts can be thought of as parts of a larger developing cancer informatics framework centered upon the evolving use of the Electronic Medical Record (EMR), EHR, and the Personal Health Record (PHR). The U.S. Department of Health and Human Services (HHS) has defined the meaning of the EMR, EHR and PHR in a report titled “Defining Key Health IT Terms”²⁶ dated April 2008, as shown in *Table 2: EMR, EHR and PHR Definitions*.

Table 2: EMR, EHR and PHR Definitions

Term	Definition
EHR	An electronic record of health-related information on an individual that conforms to nationally recognized interoperability standards and that can be created, managed, and consulted by authorized clinicians and staff across more than one healthcare organization. ²⁷
PHR	An electronic record of health-related information on an individual that conforms to nationally recognized interoperability standards and that can be drawn from multiple sources while being managed, shared, and controlled by the individual. ²⁸
EMR	An electronic record of health-related information on an individual that can be created, gathered, managed, and consulted by authorized clinicians and staff within one health care organization. ²⁹

The distinguishing difference between an EMR, EHR and PHR is that individuals control their PHRs with respect to what information is shared and with whom; authorized clinicians and staff control the EMR and EHR. An EMR is limited to the clinical care provided to a patient within a specific healthcare facility and an EHR takes a view of healthcare provided to a patient across all healthcare organizations.

These and other new technologies are dramatically changing the cancer informatics landscape. Advances in the following areas may reshape cancer care in the future:

- The next generation Internet and wireless handheld devices.
- Real-time (point of care) clinical decision support systems.
- Off-line, population-based systems (e.g., clinical reminders, clinical trials candidates, preventive screening alerts, etc.).
- Large, integrated, individual patient-level phenotypic and genotypic databases with intelligent data mining capabilities.
- Wireless, invasive and non-invasive physiologic monitoring devices.
- Natural Language Processing (NLP) systems (capable of turning free-text documents into coded clinical findings).
- Mathematical models of complex biological systems.³⁰

Cancer informatics can take on any number of combinations of objectives and tasks aimed at improving efficiency and effectiveness of cancer treatment. “Cancer information and surveillance, historically conducted with manual data collection and submission, are increasingly viewed as being inherently dependent on the effective application of information science. One of our challenges is to use information technology (IT) in a manner that improves cancer-related decision-making and ultimately the quality of care that is offered to patients with cancer.”³¹

The central focus of the NPCR-AERRO effort is to maximize the potential for cancer surveillance in light of changing technologies. To achieve this NPCR-AERRO has embarked on a series of structured intelligence gathering exercises aimed at defining needs and capabilities, current practices, areas for improvement, obstacles, and opportunities in automated cancer surveillance.

4.4 NPCR-AERRO Stakeholders

The U.S. effort to develop cancer surveillance registries and their information systems is sponsored by CDC, NCI, state governments, and professional medical associations. The current cancer surveillance programs cover 100% of the U.S. population, providing comprehensive information for research and public health assessment activities. *Table 3: NPCR-AERRO Stakeholders* lists the Cancer Surveillance stakeholders:

Table 3: NPCR-AERRO Stakeholders

Stakeholder	Representatives
Consumers/patients	
Clinicians and healthcare providers	<ul style="list-style-type: none"> • Physicians • Hospitals • Clinics • Laboratories
Payers	<ul style="list-style-type: none"> • Health Insurance Plans
Public health agencies	<ul style="list-style-type: none"> • National • State • Local
National standards-setters	<ul style="list-style-type: none"> • CDC NPCR • NCI Surveillance Epidemiology and End Results (SEER) Program • American College of Surgeons' (ACoS) Commission on Cancer (CoC) • College of American Pathologists (CAP) • North American Association of Central Cancer Registries (NAACCR)
Standards Agencies	<ul style="list-style-type: none"> • Health Level 7 (HL7) • Systematized Nomenclature of Medicine -- Clinical Terms[®] (SNOMED CT[®]) • Logical Observation Identifiers Names and Codes (LOINC[®])

Stakeholder	Representatives
Professional Organizations	<ul style="list-style-type: none"> • NAACCR • National Cancer Registrars Association (NCRA) • CAP
Software developers	<ul style="list-style-type: none"> • Hospital and Central Cancer Registry Software • EHR Software • Laboratory Information Systems • Others
Researchers	

4.5 Constraints

Limitations and restrictions to the scope of the project are described below.

4.5.1 National Standards

The cancer surveillance community has established several sets of standards for reporting cancer data, including NAACCR ePath Reporting Guidelines³² HL7³³. In addition, national organizations, such as Healthcare Information Technology Standards Panel (HITSP) and Integrating the Healthcare Enterprise (IHE), publish electronic exchange standards that dictate how data are transmitted.

4.5.1.1 Adoption

NPCR-AERRO adheres to and promotes relevant standards. However, it is outside the scope of NPCR-AERRO to establish national standards.

If enhancements to a standard are suggested as part of best practice models, the NPCR-AERRO Technical Team notes these and passes them along to the appropriate organizations for consideration. For example, additions to the NAACCR ePath Reporting Guidelines have been recommended to help streamline transmission of electronic pathology reports from laboratories to central cancer registries.

4.5.1.2 Harmonizing

Because NPCR-AERRO comprises an intersection of healthcare, public health, and technological standards, it must consider an array of standards and attempt to mesh or map data across standards as needed. For example, LOINC^{®34} and SNOMED CT^{®35} codes may need to be cross-referenced for data transmission.

Collaboration with other healthcare-focused organizations, such as the NCI and Cancer Biomedical Informatics Grid (CaBIG™), is key to across-the-board adoption of electronic reporting standards.

4.5.1.3 Licensing

Some standards, such as the CAP's Cancer Checklists and American Medical Association (AMA) Current Procedural Terminology (CPT) may require licensing.

NPCR-AERRO explores licensing, and identifies and documents issues for the cancer community to understand the impact on implementing the models.

4.5.2 EHR Adoption

EHR systems are being developed and adopted across the healthcare community.^{36, 37, 38}

NPCR-AERRO monitors and participates in EHR activities to ensure consistency with this uniform method of collecting, storing and reporting health data.

4.5.3 Consensus

NPCR-AERRO is a collaborative project among a variety of stakeholders (see *Table 3: NPCR-AERRO Stakeholders*). Consensus-building is fundamental to the model development process.

Consensus best practices are documented as the normal course of events, business rules, or system requirements in the use cases. Where differences occur (due to legal or resource constraints, for example) one or more alternate courses of events are noted.

4.5.4 Resources

Availability of resources, including funds, staff, and tools, constrains the scope and timeline of NPCR-AERRO activities and its subsequent implementation. To work within these constraints, resources are identified and provided by organizations throughout the cancer community based on the ability and willingness of stakeholders to invest in the project.

4.5.4.1 Funding

Funding of NPCR-AERRO activities is shared across the cancer community as outlined below, and may be constrained by annual budget processes and priority shifts:

- Project activities: NPCR funds the NPCR-AERRO Technical Team to develop electronic reporting models, publicize the project, and lead or participate in related activities; DCPC funds other staff time to participate in and support various aspects of the project; and cancer registry stakeholders' volunteer staff time to participate in workgroups and related activities.
- Implementation: Stakeholders fund implementation of the electronic reporting models at their organizations.

4.5.4.2 Staffing

Staffing may be constrained by the availability of appropriate types of staff to involve in the project, in addition to each organization's level of commitment to the NPCR-AERRO project based on internal priorities.

- Stakeholder organizations analyze staffing needs to manage implementation of electronic reporting models, technology upgrades, and regular operations based on the models.
- Organizations may also offer or mandate training on new technology and operations.

4.5.4.3 Tools

Tools may include software, hardware, networks, and other electronic architecture, as well as other existing cancer registry tools. Constraints are similar to other resources.

- Stakeholder organizations may require updated or new software, hardware, and/or network capabilities.
- Information architecture may be centralized or dispersed across areas of an organization, and several layers of control may be in place to approve and implement system changes.

4.5.5 Change Management of Automated Reporting Within the Cancer Registry

Any change generates resistance.³⁹ As the scope of NPCR-AERRO includes the spectrum of cancer surveillance data registration and use, it is imperative to manage the change inherent in establishment of best practice models. The NPCR-AERRO Technical Team must understand and address concerns with particular proposed changes and with the general shift toward electronic transmission of cancer surveillance data.

The NPCR-AERRO Technical Team has identified several strategies to help registries manage the changes expected through implementation of the NPCR-AERRO models:

- Training: Education on the project and specific models through web site publication, conference presentations and workshops, journal articles, and web-based training.
- Capacity Building: Identify types of resources and technical expertise needed to engage in electronic reporting.
- Technology Conservation: Note existing or cross-enterprise technology available for use in electronic cancer data reporting.
- Data and Process Integration: Diagram processes based on HL7 data exchange and electronic reporting, identify areas for operational improvement, and educate cancer registries on streamlined processes.

4.5.6 National Health Information Technology (HIT) Initiatives

NPCR-AERRO best practices should reflect national initiatives coordinated by the U.S. Department of Health and Human Services (HHS) Office of the National Coordinator (ONC), including the American Health Information Community (AHIC)⁴⁰ and the HITSP⁴¹. Requirements and standards related to the EHR should align with the goals and recommendations from these national efforts.

As new standards and guidance are issued at the national level, most healthcare systems at the hospital and physician level need significant modifications implemented to begin using required national standards. While the HHS has awarded funds to several communities, regions, states, etc. for implementation of electronic health records and standardized data exchange systems, there are not enough resources available. Funding is needed for a large proportion of the healthcare system to modify end user methodology to collect and process the data and for software vendors to implement software modifications.

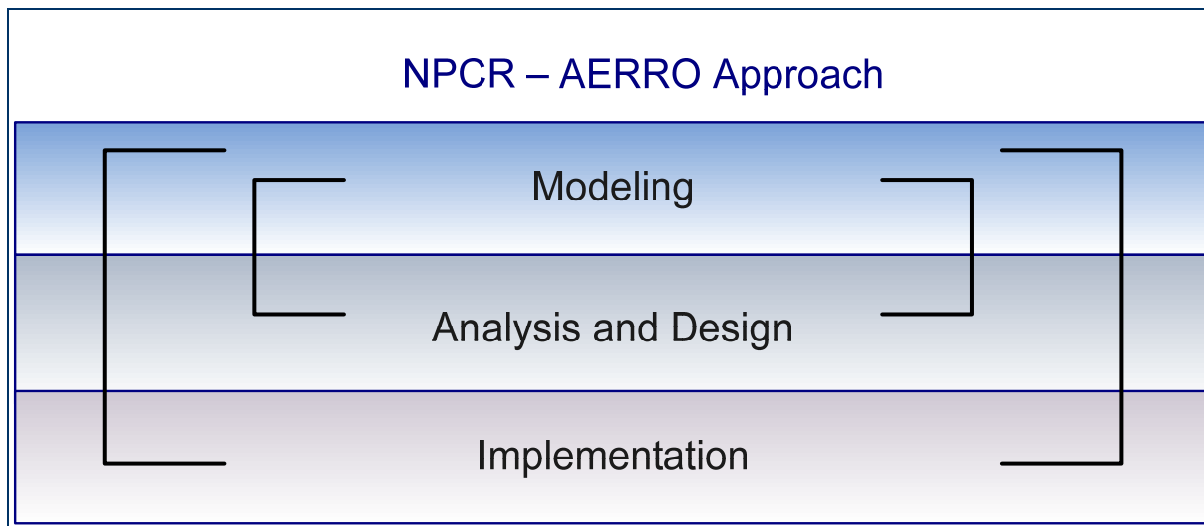
The HHS has set specific versions of HL7 as the standard for transmitting certain types of electronic health data. However, many organizations have functioning systems using prior versions of HL7 and do not have the resources nor see a need to upgrade a system that is working.

In addition, because changes in standards are at times frequent, it is in the end users' best interest to choose software and other tools that are supported and enhanced over time; the ability to make these purchasing decisions is constrained by the unknown nature of possible standards updates.

4.6 Project Approach

In order to successfully impact change in processes, the NPCR-AERRO team focuses on a three-pronged approach of Modeling, Analysis/Design and Implementation, and encompasses the full scope of cancer registration activities. As shown in *Figure 3: NPCR-AERRO Approach*, the different phases of the approach are interdependent on each other. Modeling and Analysis/Design phases overlap where the diagrams developed during Modeling phase are used as initial documents for Analysis/Design for some processes and vice versa. Outcomes of the Modeling and Analysis/Design phases of the approach feed the Implementation phase.

Figure 3: NPCR-AERRO Approach



It is useful for business processes and requirements to be modeled and documented thoroughly prior to beginning any analysis/design and implementation activities. The modeling activities fully describe the business with a specific focus on areas that could benefit from re-engineering the processes to implement automation and electronic reporting and highlight specific areas that need further analysis and design. The model must be developed by consensus at a generalized level to enable the cancer community to implement the improved standardized processes within their different respective environments. Building consensus ensures that the tools developed meet defined needs.

The analysis and design activities use products from the modeling activities to identify specific focus areas that need to be addressed. The result of the analysis and design activities is the development of detailed documentation (i.e., specific data standard, technology, or methodology) that provides additional information needed before pilot implementation can begin.

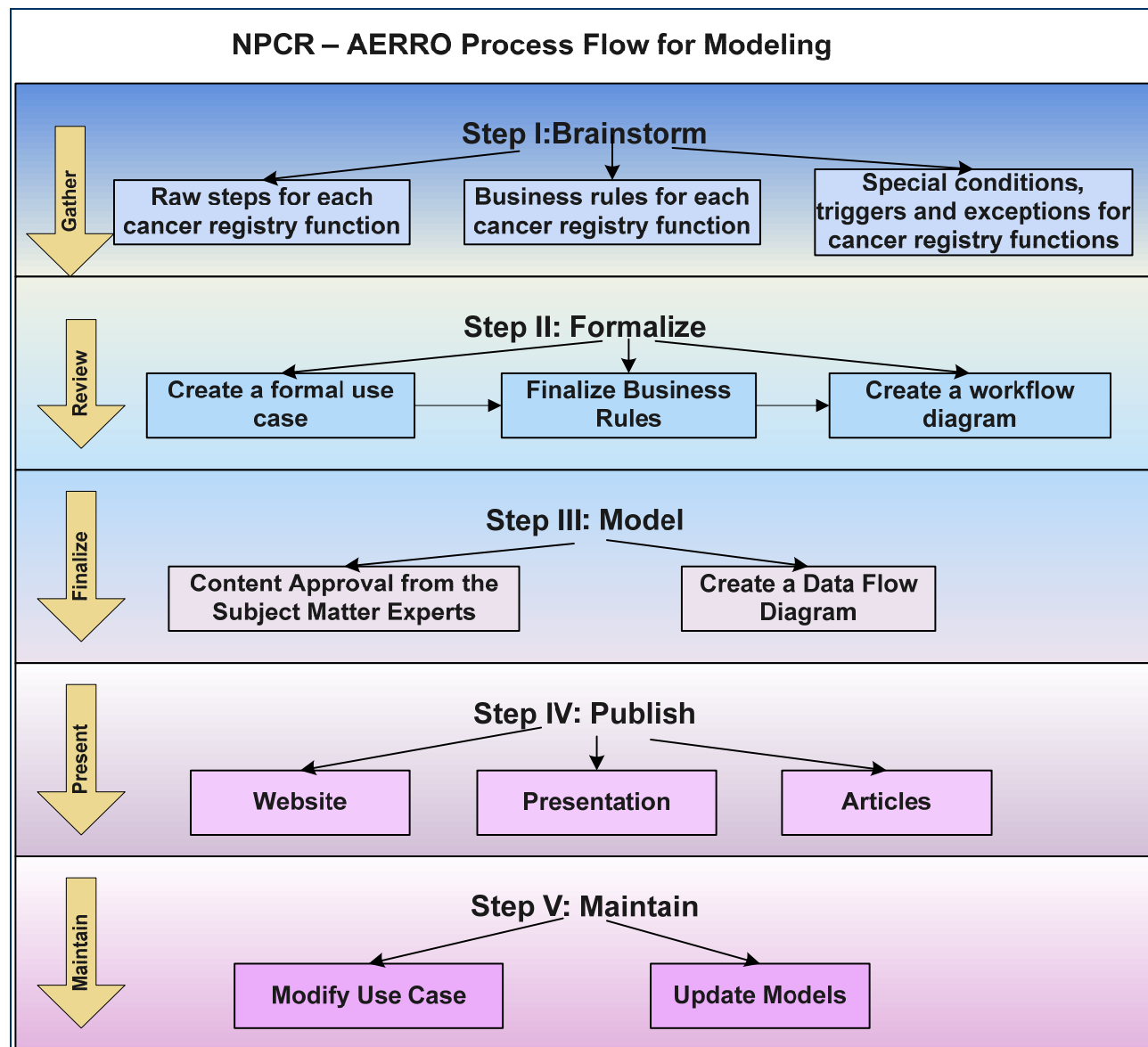
Pilot implementations should not begin until modeling and analysis/design activities are completed. This ensures that implementations address the area needed and that the solution is realistic and acceptable by the cancer community. After successful pilot implementations are complete and documented, the consensus standards and process documentation can be presented to the standard setting organizations as needed.

The next three sections describe the three approaches in more detail.

5 Modeling

NPCR-AERRO completes modeling activities to explore current business practices and to develop consensus best practice models for automating registration processes and electronic reporting. These core activities include identifying the users and developing the use cases, business rules, and Unified Modeling Language (UML) diagrams. *Figure 4: NPCR-AERRO Process Flow for Modeling* depicts the process flow used by NPCR-AERRO for conducting the modeling activities.

Figure 4: NPCR-AERRO Process Flow for Modeling



5.1 Modeling Activities

5.1.1 Strawman activity with VCU and VCR

NPCR-AERRO collaborated with the Virginia Commonwealth University Hospital System (VCUHS), the Virginia Cancer Registry (VCR), the National Cancer Institute-Surveillance, Epidemiology, and End-Results Program (NCI-SEER), and the Centers for Disease Control and Prevention National Program of Cancer Registries (CDC-NPCR) to develop a proposed “strawman” model that would be used to begin discussions with the broader cancer surveillance community. The results showed that these efforts could be broadened into a national consensus model that includes all stakeholders.⁴²

5.1.2 Strategic Assessment & Modeling Sessions (SAMS)

NPCR-AERRO conducted several SAMS to gather requirements at the Hospital and Central Cancer Registry levels as a part of Step I: Brainstorming in the “NPCR-AERRO Process Flow for Modeling diagram” (see Figure 4). A SAMS for the Hospital Cancer Registries was conducted in February, 2006 in Richmond, VA. The first SAMS for the Central Cancer Registries was conducted in March, 2006 in Atlanta, GA followed by a second SAMS held in October, 2006, again in Atlanta, GA. The participants included Certified Tumor Registrars (CTRs), data managers, statisticians, physicians, and software vendors from a number of hospitals, state and regional registries, standard-setting organizations, and software companies. During each of these requirements gathering sessions, team building exercises were carried out aimed at providing specific and detailed feedback on the following:

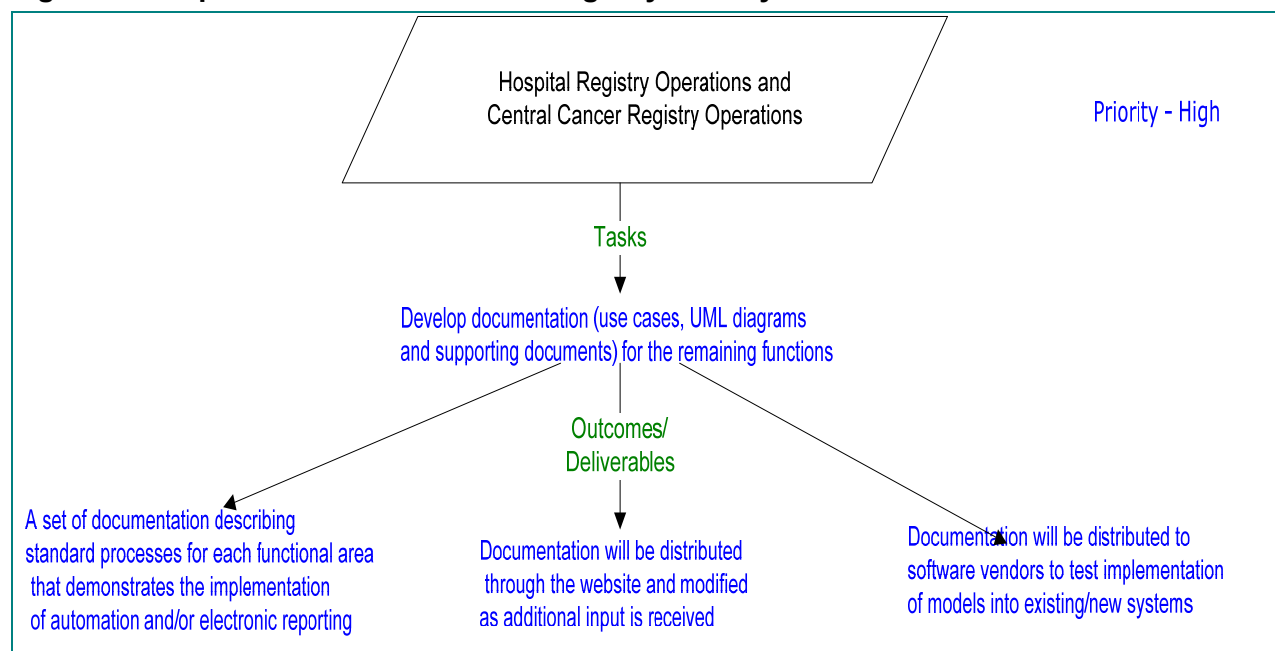
- Examples of acceptable or best practices to be duplicated in electronic reporting.
- Examples of situations and circumstances to avoid or overcome in the move toward electronic reporting.
- Barriers (technical, organizational, and content-based) needing to be addressed.
- Review of what might be possible in the next five years related to electronic reporting.
- Issues, concerns, and recommended next steps.

Each of these sessions was recorded on paper. The comments, suggestions and recommendations of the participants were later consolidated into a Requirements Finding Report⁴³ and sent to the participants. The Hospital Operations and Central Cancer Registry Workgroups were formed through these sessions.

5.1.3 Hospital and Central Cancer Registry

The objective in working with representatives from hospital and central cancer registries around the nation and associated stakeholders is to complete the domain modeling effort and to create a report of guidelines and recommendations to advance the electronic reporting initiative.

Figure 5: Hospital and Central Cancer Registry Activity Plan shows the common tasks and outcomes/deliverables for these groups.

Figure 5: Hospital and Central Cancer Registry Activity Plan

Using the SAMS information gathering process, NPCR-AERRO performed an assessment/gap analysis with each domain to determine the current state of the cancer registry business and identify priority areas for future NPCR-AERRO activity. Participants evaluated the results of the assessment/gap analysis and diagramed each domain within cancer surveillance (Appendix 1). They also defined the core functions within each domain (See Sections 5.1.3.1 and 5.1.3.2 below).

Modeling the core functions of the cancer registry occurs in monthly web-enabled conference calls with representatives from cancer registries and associated stakeholders. Workgroup members brainstorm current methods for performing a function, identify areas where electronic methods can be implemented, and develop use cases and diagrams to document a consensus best practice.

Certain consensus best practices lend themselves to further development within the modeling process. On these occasions additional documentation relating to implementing the best practice may be included as an appendix in the use case or further developed into specific implementation recommendations.

5.1.3.1 Core Functions of the Hospital Domain

Participants in the Hospital SAMS identified four core functions and several associated functions in hospital registry operations, shown in *Table 4: Hospital Cancer Registry Core Functions*.

Table 4: Hospital Cancer Registry Core Functions

Hospital Cancer Registry Core Functions	
<ul style="list-style-type: none"> • Creating a Cancer Registry Record <ul style="list-style-type: none"> ○ Data source preparing an Event Report for availability to the hospital cancer registry ○ Receiving, validating and determining the most accurate and complete 	

Hospital Cancer Registry Core Functions	
data reflected in the multiple Event Reports	
○	Casefinding
○	Abstracting
○	Editing
•	Reporting
•	Enhancing the Data
○	Quality Assurance/Quality Control
○	Follow-Up (as required by the standard setters)
•	Analysis

The core functions are translated into 10 use cases as depicted in the Hospital Registry Operations Use Case Diagram (Appendix 2).

Some functions, such as reporting cases to the central registry, have been effectively automated. Other functions, such as editing, currently have a mix of manual and electronic methods. NPCR-AERRO evaluates if and how electronic methods can replace manual methods and how to efficiently integrate them with existing electronic processes.

It is in the areas of casefinding, abstracting and follow-up that AERRO can make a major impact on hospital cancer registry operations. A focused effort on casefinding can significantly improve the timeliness and completeness of cancer reporting by identifying cancer cases more quickly and accurately. A focused effort on abstracting, such as automatically inserting event report data into the cancer registry abstract with both abstract and event report available in the facility's information system, has several advantages:

- Reduces transcription errors, providing more accurate data.
- Decreases the amount of time used for collecting the data.
- Allows the opportunity to collect additional data without additional resources.

Moving the follow-up process into an electronic environment can reduce the amount of staff time spend on most of the clerical activities relating to follow-up activities, allowing staff to spend more time on the higher level tasks. *Table 5: Potential Impact of NPCR-AERRO Model* describes how implementation of the NPCR-AERRO business model might impact a hospital registry.

Table 5: Potential Impact of NPCR-AERRO Model

Impact of Implementing the NPCR-AERRO Model in a Hospital Cancer Registry	
Improve timeliness, completeness and validity of many data elements	
•	Automate flow of case finding
•	Potential for real time case ascertainment
Capture additional important data elements without increasing workload	

Impact of Implementing the NPCR-AERRO Model in a Hospital Cancer Registry
<ul style="list-style-type: none"> • Follow up information
<ul style="list-style-type: none"> • Clinical diagnostic parameters (e.g., tumor markers, mitotic rate, etc.)
<ul style="list-style-type: none"> • Recurrence—Date of recurrence and type
Enhance the job duties and activities of the registrar
<ul style="list-style-type: none"> • Provide an opportunity for increasing use of registry specialized analytic skills and activities
<ul style="list-style-type: none"> • Increase computer knowledge and technical skills
<ul style="list-style-type: none"> • Diminish clerical/paper work

5.1.3.2 Core Functions of the Central Cancer Registry Domain

Central cancer registry stakeholders identified seven major core functions and several associated functions in registry operations, as shown in *Table 6: Central Cancer Registry Core Functions*.

Table 6: Central Cancer Registry Core Functions

Central Cancer Registry Core Functions
<ul style="list-style-type: none"> • Creating a Cancer Registry Record <ul style="list-style-type: none"> ○ Data source preparing an Event Report for submission to the central cancer registry ○ Receiving and validating the Event Reports ○ Performing patient linkage, tumor linkage, and record consolidation
<ul style="list-style-type: none"> • Enhancing Data <ul style="list-style-type: none"> ○ Performing audits, quality assurance and quality control ○ Performing external linkages to improve the data ○ Conducting death clearance and follow-up
<ul style="list-style-type: none"> • Data Exchange with other Registries
<ul style="list-style-type: none"> • Calls for Data
<ul style="list-style-type: none"> • Using the Data <ul style="list-style-type: none"> ○ Disseminating data for use by others ○ Conducting Linkage for Research ○ Performing analysis and research
<ul style="list-style-type: none"> • Education and Training

Central Cancer Registry Core Functions

- Security

The core functions are translated into 20 use cases as depicted in the Central Cancer Registry Operations Use Case Diagram (Appendix 2).

Even more than in the hospital domain, many central cancer registry functions, such as receiving and validating cancer registry event reports, already exist as electronic functions. NPCR-AERRO may contribute only minimally to enhancing those functions. Within the same function, however, receiving and validating event reports from data sources other than the hospital is less well defined. NPCR-AERRO evaluates if and how electronic methods can be efficiently integrated into existing processes.

For more information please visit the project site on the internet:

<http://www.cdc.gov/cancer/npcr/informatics/AERRO/workgroups/central>.

5.1.4 National Cancer Programs

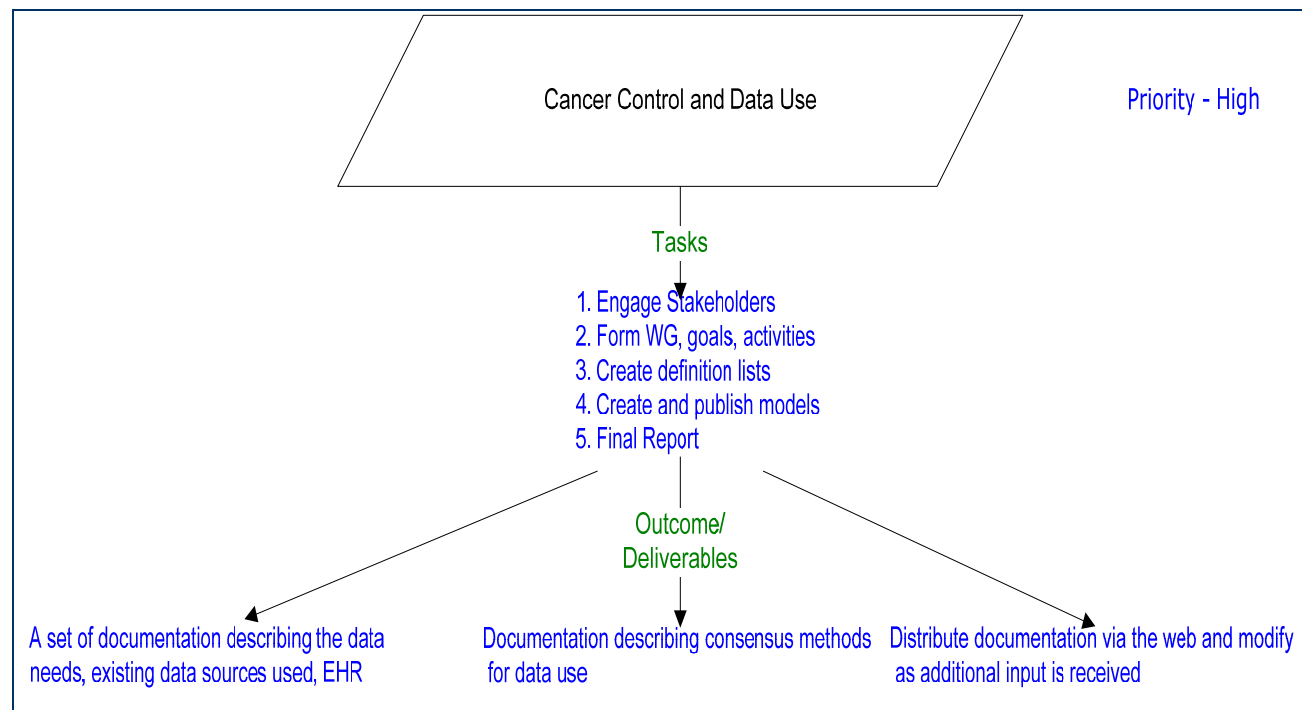
The activities of NPCR-AERRO at the hospital and central cancer registry level can have a positive impact on the national programs' ability disseminate cancer data in a more timely manner. The activities also improve the already high quality and completeness of the data.

NPCR-AERRO explores the feasibility of modeling business processes of the national cancer programs, especially in the area of producing cancer publications and data files. Use cases from the central cancer registries may be leveraged for this activity.

5.1.5 Cancer Control and Data Use

The NPCR-AERRO Cancer Control and Data Use (CC&DU) initiative evaluates how electronic reporting of cancer data and adoption of the Electronic Health Record (EHR) can impact use of cancer surveillance data, such as improved timeliness and data quality for better trend analysis. This forum will consist of domain experts such as scientists, administrators, program managers, health plan representatives, and technology professionals.

Figure 6: Cancer Control and Data Use Activity Plan shows the tasks and outcomes/deliverables for the workgroup.

Figure 6: Cancer Control and Data Use Activity Plan

The goals of the Cancer Control and Data Use initiative are to:

- Model how cancer data is used for current and future needs.
- Identify data sources and/or data elements not previously available.
- Model feedback mechanisms between data use and cancer surveillance.

The core activities consist of:

- Examining the short- and long-term data needs of cancer surveillance data users.
- Identifying how cancer surveillance data are used to inform health and administrative decision making.
- Classifying current and potential users of cancer registry data.
- Identifying efficient patterns of usage that can model as best practices.
- Maximizing the electronic infrastructures of EHRs and Patient Health Records (PHRs) in collecting and distributing cancer-related statistical data.
- Defining knowledge products that can be enhanced by electronic cancer data exchange.

5.1.6 e-Health Initiatives

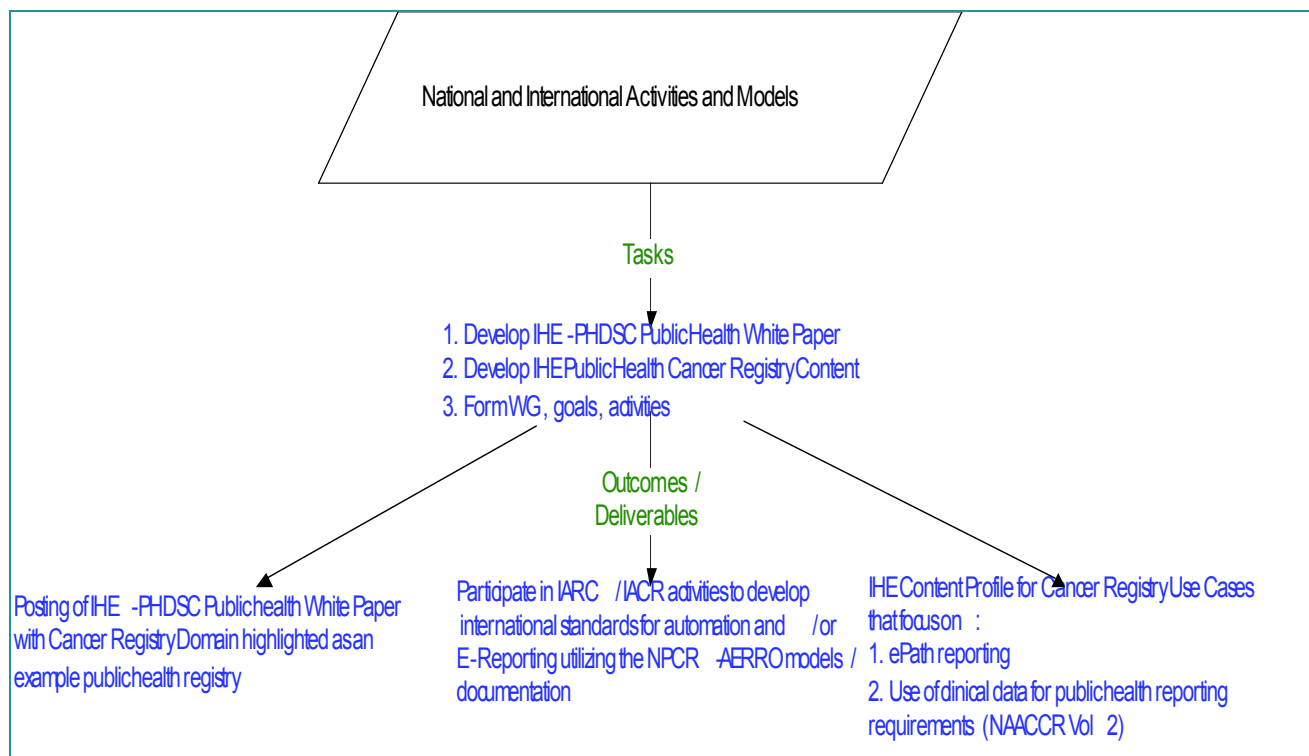
In an effort to demonstrate how public health fits into the implementation of the EHR, NPCR-AERRO collaborated on the Public Health Data Standards Consortium's (PHDSC) White paper, "Building a Roadmap for Health Information Systems Interoperability for Public Health". The overall goal of this effort is to facilitate the necessary linkages, standardization, and integration of health data between clinical care and public health to create robust overarching health information exchanges. The objective is to engage the public health community in a dialogue with health information technology (HIT) vendors to assure that the work processes and data needs of public health stakeholders in health information exchanges are 1) well understood and

agreed upon by stakeholders themselves, and then 2) communicated clearly to the developers of the interoperable clinical EHR systems and Public Health information systems (EHR-PH Systems).⁴⁴ The white paper uses the cancer surveillance community as an example public health domain.

More specifically NPCR-AERRO models how the cancer surveillance community interfaces with the national e-Health initiatives. This model can inform the community on regulations, guidelines and best practices that are being disseminated that have an impact on their processes and objectives.

Figure 7: National and International Activity Plan shows the tasks and outcomes/deliverables for these activities.

Figure 7: National and International Activity Plan



5.1.7 Implementation of Electronic Reporting for Data Sources

NPCR-AERRO plans to work with data sources that submit information to central cancer registries to implement electronic reporting using standardized methodologies.

5.1.7.1 Physician Offices/Oncology Clinics

As more and more cancers are diagnosed and/or treated in the physician's office setting, it is important for the central cancer registries and the physician offices/cancer clinics to understand the processes that need to be implemented in order to have complete, timely, and accurate cancer case reporting.^{45,46} Physician office/cancer clinic specialties need to be fully described and understood to identify methods for automated and electronic reporting. NPCR-AERRO will develop models and associated documentation to describe the different types of physician offices/cancer clinics that would provide data to the cancer registries and the processes used to

do so. The documentation will describe the types of data needed and develop triggers that will assist the physician/clinic with identifying specific cases as reportable, such as a decision support system. Throughout this process, NPCR-AERRO will include stakeholders from the broad cancer community as well as the physician offices, cancer clinics, and software vendors.

For more information please visit the project website on the internet:

<http://www.cdc.gov/cancer/npcr/informatics/AERRO/workgroups/data.htm>.

5.2 Technical Approach for Modeling

5.2.1 Introduction

The methods of exchanging and distributing data have undergone a major change in the past 10 years to take advantage of advances in technology. NPCR-AERRO makes use of the best practices from related projects worldwide in order to not duplicate efforts. NPCR-AERRO uses frameworks and methodologies which focus on reuse of existing processes and an iterative development approach.

5.2.2 The Use of Modeling in Cancer Registration

Cancer registration is a complex process by which the registries collect, consolidate, enhance, maintain, and report data to the multiple state, national, and international organizations. Because of the different stakeholders involved in the area of cancer registration, building consensus is important. Modeling facilitates communication among stakeholders, and promotes analysis and improvements in the essential aspects of cancer registration.⁴⁷ Also, visual modeling can help identify the gaps in the number of existing electronic software systems which can help registries improve and update their current systems.

The cancer registry community has used modeling for a number of years to help describe and document the complexities of its processes, its databases and the relationships between its participants. Modeling projects include:

5.2.2.1 NCICB's caCORE⁴⁸

caCORE is open source software and services developed by the National Cancer Institute Center for Bioinformatics (NCICB) Core Infrastructure Group. Two key components of caCORE are the Enterprise Vocabulary Services (EVS) and the Cancer Data Standards Repository (caDSR). An Iterative Software Development Approach has been used for caDSR. The Iterative Software Development Approach is a combination of both Rational Unified Process (RUP) and eXtreme Programming (XP).⁴⁹ The use case model has helped ensure that all the right features for the stakeholders have been included and makes it easy to see the whole system. They are documented simple text or UML diagrams drawn by hand or with a modeling tool such as Rational Rose. Use cases are prioritized based upon several factors such as risk, importance to the customer, and technical difficulty.⁵⁰

5.2.2.2 NCI's SEER*DMS, 2000

SEER Data Management System (SEER*DMS) was a project started in 2000 by NCI's SEER program to collect cancer incidence and related information from population-based cancer registries automatically. It provides support for all core cancer registry functions—importing data, editing, linkage, consolidation, and reporting.⁵¹ The system design and models were based on Chen's Entity-Relationship Model and the methodology used was that of Matt Flavin.⁵² Joint Application Development (JAD) sessions were conducted to gather requirements

with a number of stakeholders such as NCI, SEER Registries, State Health Registries, vendors and software developers.

5.2.2.3 NAACCR E-Path Modeling Project

In 2006, NAACCR released the Electronic Pathology (E-Path) Reporting Guidelines⁵³ to define the recommended approach for implementing standards for electronic pathology reporting. UML diagrams and use cases were used to describe the process in a way that is understandable by end users and IT personnel.

5.2.2.4 Central Cancer Registry Modeling Project, 1999

This was a collaborative effort started to identify the differences and similarities that exist between the cancer registry system and other health information and disease surveillance systems available through modeling. CDC-NPCR, NCI-SEER, State Cancer Registries and Software Vendors were the major stakeholders. The use of modeling helped the stakeholders to analyze every aspect of existing systems in a structured way and to identify modifications which would improve and enhance these systems.

5.2.3 The Modeling Strategies and Techniques

The following are technical strategies that can and have been used by NPCR-AERRO and other projects:

- Entity-Relationship Modeling – This technique was developed by Dr. Peter Chen⁵⁴ in 1976. Entity-relationship modeling is a relational schema database modeling method used in software engineering to produce a type of conceptual data model (or semantic data model) of a system, often a relational database, and its requirements in a top-down fashion. Diagrams created using this process are called *entity-relationship diagrams*, or *ER diagrams* or *ERDs* for short.⁵⁵
- Agile Modeling – The originator of agile modeling is Scott W. Ambler. It is a practice-based methodology for effective modeling and documentation of software-based systems. Simply put, Agile Modeling is a collection of values, principles, and practices for modeling software that can be applied on a software development project in an effective and light-weight manner.⁵⁶
- IBM's Rational Unified Process⁵⁷ – The Rational Unified Process (RUP) is a Software Engineering Process. It provides a disciplined approach to assigning tasks and responsibilities within a development organization. Its goal is to ensure the production of high-quality software that meets the needs of its end-users, within a predictable schedule and budget. The RUP provides each team member with the guidelines, templates and tool mentors necessary for the entire team to take full advantage of best practices, including:
 - Develop software iteratively.
 - Manage requirements.
 - Use component-based architecture.
 - Visually model software.
 - Verify software quality.
 - Control changes to software.

RUP is a guide for how to effectively use the UML. The UML is an industry-standard language that allows people to clearly communicate requirements, architectures and designs. It helps specify, visualize, and document models of software systems and non-software systems, including their structure and design at a higher level of abstraction by bringing out the big picture.⁵⁸ The UML was originally created by Rational Software, and is now maintained by the standards organization Object Management Group (OMG).

- CDC Unified Process - The CDC Unified Process (UP) is a methodology developed from IBM's RUP. It is a collection of processes, tools, and artifacts that any project can use to structure, track, and manage their activities and deliverables. The CDC UP is a defined and clear approach to successful project delivery through a consistent and repeatable integration of practices and processes that comply with Federal regulations and policies, industry best practices, and Public Health Information Network (PHIN) and CDC standards.⁵⁹

5.2.4 NPCR-AERRO Modeling Plan & Use Case Development

NPCR-AERRO is a step towards automation of cancer registries at the (1) hospital level (2) state-wide population based central cancer registries, and (3) national cancer programs (See *Figure 1: NPCR-AERRO Context Diagram*). NPCR-AERRO is using modeling to visually represent the processes for reaching a consensus and identifying best practices for cancer registration. Business processes for each core function at each level are being documented textually and visually by NPCR-AERRO.

Because of the suitability of the CDC UP and UML, these two techniques are being used by the NPCR-AERRO Technical Team to plan, gather requirements, analyze and design its results in the form of use cases, models, business rules, software requirements and other useful reference documents. A definition of use case, model, business rules and software requirements is provided in Section 5.2.6 below.

5.2.5 Use of UML

As mentioned above, UML is a formal modeling language which is helpful in creating models that are robust, scalable and transportable, so that they can be used to develop customized software for different reporting facilities with little or no modification. UML is not simply a notation for drawing diagrams, but a complete language for capturing knowledge (semantics) about a subject and expressing knowledge (syntax) regarding the subject for the purpose of communication⁶⁰. It has the ability to combine principles, techniques, methods and tools. Through visual modeling and UML, almost anything can be modeled to be refined and modified later during the software development phase. NPCR-AERRO uses IBM's Rational Rose as the software tool to develop the models.

Some benefits of UML⁶¹ are:

- Systems are professionally designed and documented before they are coded so that all stakeholders know exactly what they are getting, in advance.
- UML enables logic 'holes' to be spotted in design drawings so that the system and software behave as expected.
- Since system design comes first, UML enables re-usable code to be easily identified and coded with the highest efficiency, thus reducing software development costs.
- The overall system design described in UML dictates the way the software is developed so that the right decisions are made early on in the process. Again, this reduces software development costs by eliminating re-work in all areas of the life cycle.

Some limitations of UML are:

- Unified Modeling Language can be cumbersome and complex due to the number of different diagram types available. These diagram types can be redundant and repetitive.
- It may be difficult to learn and understand and can be considered aesthetically inconsistent because of the mix of ovals and boxes.
- Although UML gives proper de-notational code, it does not give code that can be converted to a program immediately and so it is non-executable⁶²; that is, the code given by UML needs additional work to be converted to a functional program.

5.2.6 UML Concepts used in NPCR-AERRO

There are numerous concepts of UML available which can be used for project design and development. Some of the main UML concepts used by NPCR-AERRO are defined in this section. They are:

Use Case: The specification of sequences of actions, including variant sequences and error sequences, that a system, subsystem, or class can perform by interacting with outside objects to provide a service of value is called a Use Case.⁶³

Model: A model plays the analogous role in software development that blueprints and other plans (site maps, elevations, physical models) play in the building of a skyscraper.⁶⁴

Business Processes: A description of a set of related activities that, when correctly performed, satisfy an explicit business goal.

Business Rules: Statements that constrain, derive, and give conditions of existence are called business rules. Business Rules are used to specify allowed state of affairs, including allowed business object states.

Software Requirements: Statements that constrain the design and implementation of a software application are called software requirements.

NPCR-AERRO makes use of the diagrams shown in *Table 7: NPCR-AERRO Diagrams*.

Table 7: NPCR-AERRO Diagrams

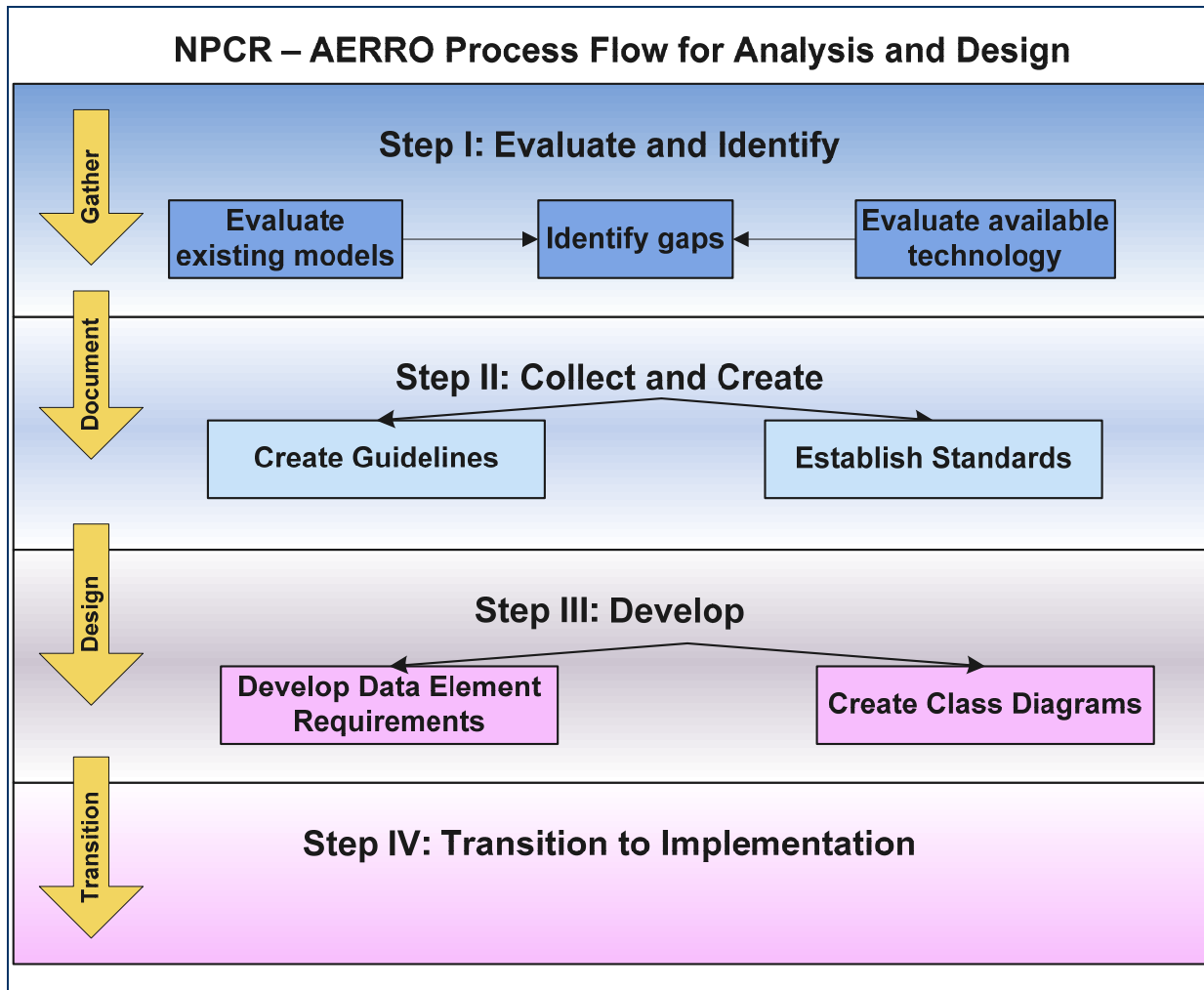
Types of Model	Definition	Sample AERRO Diagrams
Use Case Diagram <ul style="list-style-type: none"> • Operations Use Case Diagram • Business Use Case Diagram 	Presents a high level view of how the system is used from a user's perspective	<ul style="list-style-type: none"> • CCR Operations Use Case Diagram • CCR Business Use Case Diagram
Class (Domain) Diagram	Depicts high-level units of possible systems.	<ul style="list-style-type: none"> • CCR Domain Diagram
Activity Diagram <ul style="list-style-type: none"> • Workflow Diagram • Data Flow Diagram 	Provides a way to model the workflow of a business process. These are similar to flowcharts as a workflow can be modeled from activity to activity.	<ul style="list-style-type: none"> • CCR Receive Batch File Workflow Diagram • CCR Receive Batch File Data Flow Diagram

Examples can be found in Appendix 1 and Appendix 2.

6 Analysis and Design

NPCR-AERRO analyzes current technology and infrastructure surrounding registry operations based on the NPCR-AERRO models and designs specifications for the development of products to support automation and electronic reporting. Results of analysis and design may lead to functional specifications or feature enhancements for registry software, implementation guides, class diagrams, white papers or gap analyses, and identification of standards. *Figure 8: NPCR-AERRO Process Flow for Analysis and Design* depicts the process flow used by NPCR-AERRO for conducting the analysis and design activities.

Figure 8: NPCR-AERRO Process Flow for Analysis and Design



6.1 Analysis and Design Activities

6.1.1 Hospital and Central Cancer Registry (CCR) Functions

6.1.1.1 Visual Editing

Visual editing looks for logical consistency among data fields and verifies by reviewing the supporting documentation. There are specific data items that are critical for central registry use

and therefore, lend themselves to visual editing. The number of cases and number of data items included in the visual editing process should not outweigh the cost/benefit. Using a standardized representative sampling method should improve the validity of the visual editing process as it provides a more statistically stable number of cases than a 1:10 or 1:25 ratio currently being applied to a batch of abstracts. The NPCR-AERRO has developed a sampling method for registries to set an appropriate number of cases to visually edit as a routine quality assurance process.

6.1.1.2 Feedback Form from the CCR to the Hospital Cancer Registry

NPCR-AERRO is evaluating and designing a template for CCRs to use to provide feedback to hospital cancer registries that submit cancer data. Providing information regarding the central registry's findings when processing the file can help hospital registries identify problem areas and improve their submissions.

NPCR-AERRO will explore developing a software module that the CCRs can use to create a feedback form.

6.1.2 Electronic Pathology Reporting

6.1.2.1 Text Mining/Natural Language Processing for Mapper Plus

NPCR-AERRO has evaluated several open source natural language processing software packages for possible inclusion in Mapper Plus, a software program that processes HL7 messages. (See Section 7.1.2.2 for a full discussion of Mapper Plus.) Further analysis may identify areas where Mapper Plus' natural language processing function can be improved to increase the accuracy in selecting cancer cases.

6.1.2.2 Mappings of the College of American Pathologists electronic Cancer Checklist (CAP eCC)

The CAP Cancer Committee "publishes cancer checklists to assist surgical pathologists in reporting both common and uncommon forms of cancer". These checklists have been encoded with SNOMED CT codes and are available as electronic tools for use in the anatomical pathology community. Collaborative efforts are underway to map the CAP eCC concepts to the NAACCR data items, including those related to Collaborative Stage.⁶⁵

6.1.2.3 Collaboration with Cancer Surveillance Standard Setters

NPCR-AERRO is committed to collaborating with cancer surveillance stakeholders to ensure a single consistent electronic pathology (ePath) process can be developed and maintained. There is cross-participation on committees and workgroups to analyze the workflow and data requirements and design standard guidelines and procedures for ePath reporting.

6.1.3 National E-Health Initiatives

NPCR-AERRO collaborates with the Healthcare Information and Management Systems Society (HIMSS) Integrating the Healthcare Enterprise (IHE) domains to test a Cancer Registry Pathology Reporting Profile that would transmit information from pathology laboratories to the appropriate state cancer registries using the NAACCR Standards for Cancer Registries Volume V: Pathology Electronic Reporting HL7 v. 2.3.1 ORU message format.⁶⁶ This is the use case that the cancer registry community proposed for inclusion in an international framework. This activity brings the data needs of the cancer registry to the attention of EHR software developers,

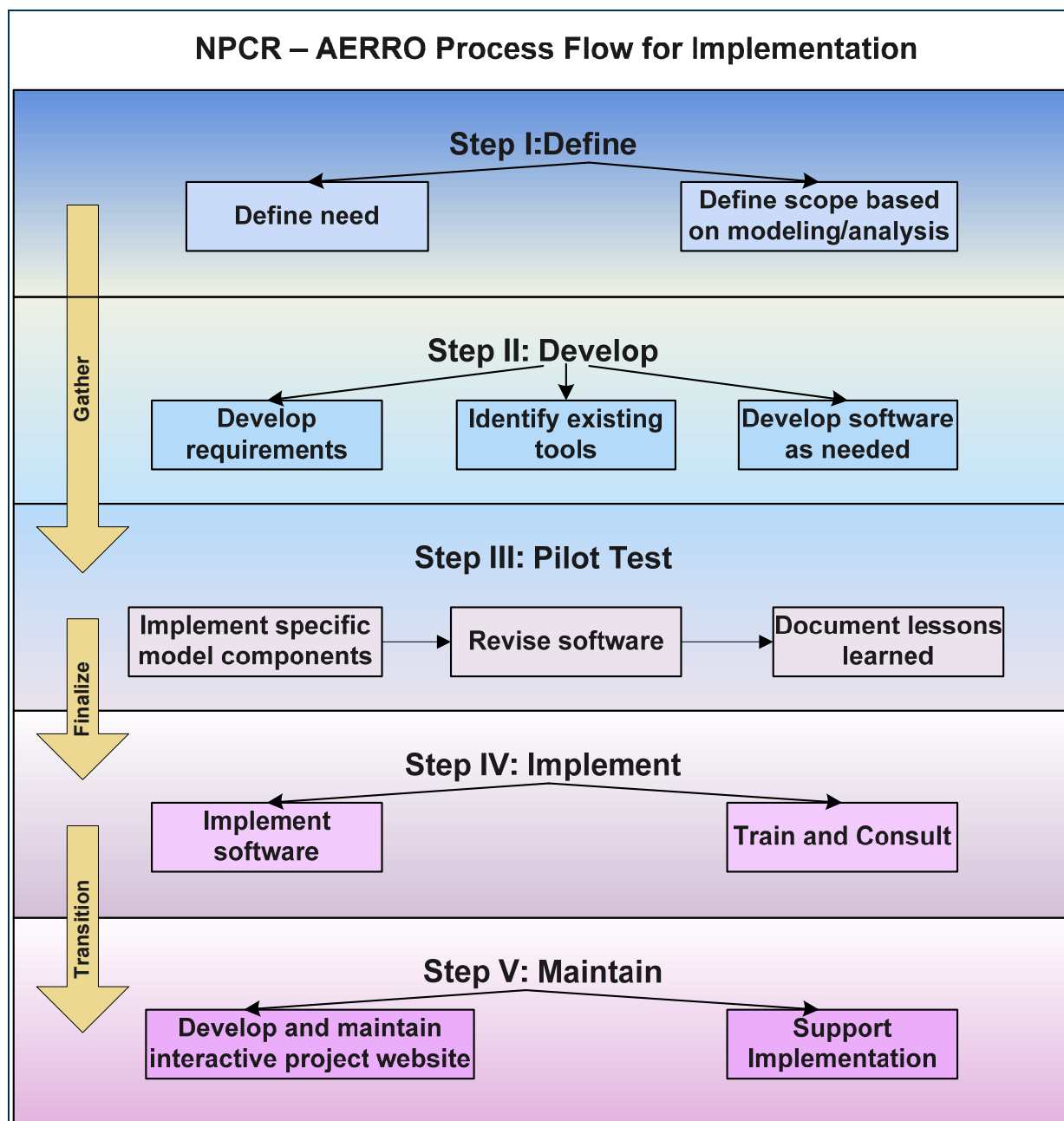
thus ensuring that cancer registry data needs are included as a national standard for EHRs as they are developed.

NPCR-AERRO will evaluate additional use cases for inclusion in the IHE framework, including use of summarized cancer registry data to inform clinician/patient interactions.

7 Implementation

NPCR-AERRO participates in implementation core activities by coordinating, leading, and supporting software vendors, hospitals, and state cancer registries to pilot test the NPCR-AERRO models and products. *Figure 9: NPCR-AERRO Process Flow for Implementation* depicts the process flow used by NPCR-AERRO for conducting the implementation activities.

Figure 9: NPCR-AERRO Process Flow for Implementation



7.1 Implementation Activities

7.1.1 Hospital and Cancer Registry Functions

7.1.1.1 Multiple Primary Determination Module for NPCR Registry Plus

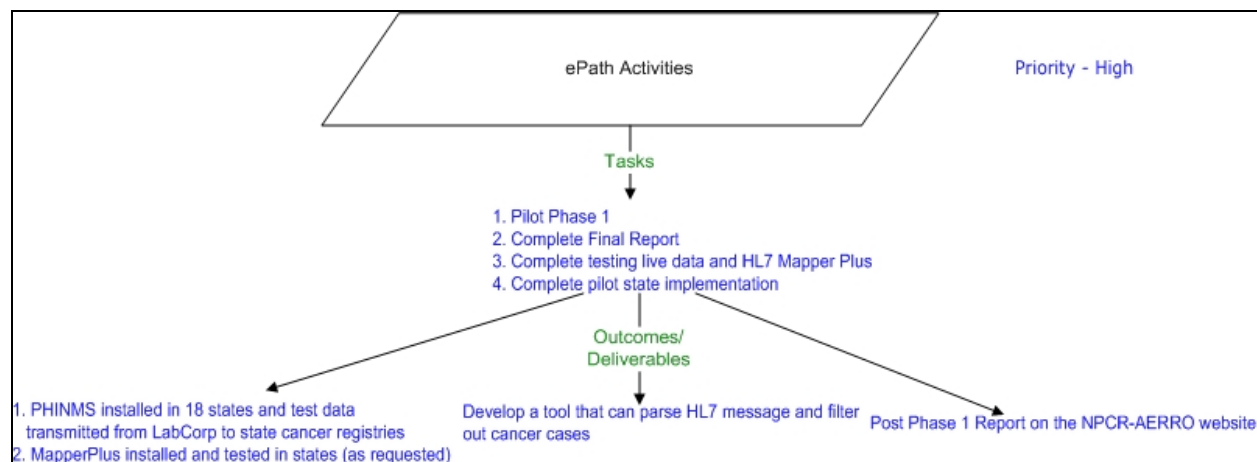
NPCR-AERRO identified “performing tumor linkage” as a use case within the central cancer registry. In this use case, event reports that have been received for the same patient are evaluated to determine whether the reports represent a single cancer or more than one cancer. The evaluation is most often performed manually, requiring staff resources and causing delays in the availability of finalized cancer information. Working with the Florida Cancer Data System and the NPCR Registry Plus Development Team, NPCR-AERRO is developing a software module to automate tumor linkage for an estimated 80 and 95% of cancer cases.

7.1.2 Electronic Pathology Reporting

The information collected and included in the pathology laboratory reports represents a critical data source for state cancer registries. Currently, some registries still lack the resources either to obtain and process paper pathology reports or to implement their own electronic pathology (ePath) reporting systems. Many laboratories lack resources and infrastructure to implement ePath reporting. The need to retrieve data from the pathology report in a more efficient and timely fashion is driving the development of an automated electronic process for accessing and utilizing pathology reports to identify cancer cases.

Figure 10: Electronic Pathology Reporting Activity Plan shows the tasks and outcomes/deliverables of this activity.

Figure 10: Electronic Pathology Reporting Activity Plan



7.1.2.1 ePath Pilot Project

NPCR-AERRO is conducting a pilot project to test the implementation of transmitting electronic anatomical pathology reports from a national laboratory to state central cancer registries. This pilot project has the potential to move the cancer registry community forward in using consistent standards for ePath reporting that can improve the completeness, timeliness, and quality of cancer registry data.

standards for ePath reporting that can improve the completeness, timeliness, and quality of cancer registry data.

The ePath Pilot Project was formed as a collaboration between Centers for Disease Control and Prevention's (CDC) NPCR-AERRO, Laboratory Corporation of America® Holdings (LabCorp®), CDC's Public Health Information Network (PHIN), and 18 state central cancer registries including Alabama, Arizona, California, Colorado, Florida, Georgia, Michigan, Missouri, Nevada, New Hampshire, New Jersey, New York, North Carolina, Ohio, Oklahoma, Tennessee, Texas, and Virginia.

This pilot project tests electronic anatomical pathology reporting from LabCorp®, a national laboratory, to state central cancer registries using the recently approved standard in the North American Association of Central Cancer Registries (NAACCR) Volume V Standard for Pathology Laboratory Electronic Reporting and the business rules defined in the draft NAACCR ePath Reporting Process Guide⁶⁷. It demonstrates the ability to integrate electronic pathology reporting for cancer registries into the infrastructure for electronic laboratory reporting of communicable diseases and bio-surveillance.

Phase I of the ePath Pilot Project was completed in 2007; a final report is published on the internet in both PDF and Microsoft Word formats:

http://www.cdc.gov/cancer/npcr/npcrpdfs/AERRO_epath_final_2007.pdf

http://www.cdc.gov/cancer/npcr/npcrpdfs/AERRO_epath_final_2007.doc.

The ePath Pilot Project is transitioning to the ePath Implementation Project, collaborating with other national laboratories to initiate electronic reporting of cancer cases. NPCR-AERRO has begun working with Bostwick Laboratories, U.S. Labs, Dianon Laboratories, Mayo Medical Laboratories, and Quest Diagnostics. Initial work with these labs has included providing an orientation of the requirements for implementing electronic reporting using the NAACCR Volume V standard and the Public Health Information Network Messaging System (PHINMS); and providing guidance on the development of the Health Level 7 (HL7) v2.3.1 ORU message for testing and validation.

7.1.2.2 Mapper Plus

Mapper Plus, the newest of the Registry Plus applications, is an application developed to view and work with HL7 files and messages. The application is developed collaboratively by participants in the NPCR-AERRO's ePath Pilot Project and programmed by the Registry Plus Development Team.

The program includes functions to import HL7 files manually or directly from the PHINMS queue, test messages for existence of required data items, parse HL7 messages, and map HL7 data elements to NAACCR data elements. Mapper Plus also builds a pathology lab database, storing various HL7 data elements as discrete field values into tables in the database. The program creates NAACCR formatted abstract records from pathology reports during import into the pathology lab database. The program searches a terms table to find potential reports of cancer, and the negation terms finder algorithm (NegEx) has been built to enhance the program's text mining capabilities in terms of specificity. Mapper Plus provides a screen to view pathology report text and a generated abstract side-by-side to allow coding of primary site and histology; it allows users to override any automated decisions about reportability and coding. Further development has been planned to identify site and morphology codes or provide a list of codes to users as they work with pathology report-generated abstracts in this program.

For more information including future project activities, please visit the project website on the internet: <http://www.cdc.gov/cancer/npcr/informatics/AERRO/workgroups/epath.htm>.

7.1.3 e-Health Initiatives

NPCR-AERRO's participation in the Integrating the Healthcare Enterprise (IHE) Pathology Domain will culminate in the implementation of the accepted profile into one or more vendors' Electronic Health Record (EHR) software applications. The ability to electronic report cancer information using the EHR can reduce the resources needed for a hospital or pathology laboratory to meet their cancer reporting obligations.

8 Project Management

8.1 Methodology

NPCR-AERRO project management is based on the Centers for Disease Control and Prevention Unified Process (CDC UP) which provides a framework with processes and templates to support best practices in project management.⁶⁸

Using the CDC UP framework, NPCR-AERRO project management is structured around development phases and releases:

Development phases: Lifecycle of creating best practice models, requirements, and implementation pilots for electronic transmission of cancer data.

Releases: Publication for community use after completion, approval, and clearance.

The overall focus for NPCR-AERRO project management is quality management based on consensus-building among identified stakeholders, delivered according to agreed-upon scope and timelines.

8.2 Development Phases

CDC UP includes the following development phases⁶⁹ which form the structure of NPCR-AERRO project management. These phases apply to the overall NPCR-AERRO project and to major activities within the project, such as hospital operations and electronic pathology reporting.

8.2.1 Initiation

The Initiation phase documents the vision, high-level scope, and stakeholders for the project.

Focus for initiating the NPCR-AERRO project:

- **Variety of stakeholders:** Because NPCR-AERRO is consensus-based, identifying representative cancer community members is key to producing comprehensive and useful products.
- **Scope definition:** The numerous possible activities involved with electronic transmission of cancer data require defining the domain of NPCR-AERRO, determining in- and out-of-scope activities, and prioritizing in-scope activities.

8.2.2 Planning

The Planning phase elaborates project requirements:

- Scope: Deliverables, timeline, stakeholder roles and responsibilities, plus ideas for future consideration
- Project plan: Work breakdown structure, leads, timelines, dependencies
- Communications plan: Internal and external channels for status and publications, document control, and update frequency
- Change control plan: Scope management with impact to deliverables and timelines
- Risk management plan: Possible issues which likely impact and mitigation strategies

- Quality control plan: Criteria and timelines for acceptance of project deliverables

Focus for planning the NPCR-AERRO project:

- **Scope of work:** Many project activities overlap or are dependent on other activities, so careful definition of scope for each activity is essential.
- **Communication:** Communication channels within the project team, among stakeholders, and to the cancer community must account for remote locations, time differences, and access limits.

8.2.3 Executing

The Executing phase focuses on product creation, review, revision, and approval.

Focus for executing models and supporting materials for the NPCR-AERRO project:

- **Activities:** Planning and execution of major activities within NPCR-AERRO may be independent or synchronized depending on scope dependencies; for example, Central Cancer Registry and Hospital Operations can be modeled concurrently, but may require some joint workgroup sessions for cross-domain use cases.
- **Products:** The main output of NPCR-AERRO consists of best practice models, business and system requirements, and implementation pilots.

8.2.4 Closing

Project Closing ends the lifecycle of project development for a release, including the final timeline, reports, publications, document control, and suggestions for future development.

Focus for closing out NPCR-AERRO project activities:

- **Activity closeout:** Each release of a major activity within NPCR-AERRO is closed out according to its particular requirements and timeline.
- **Maintenance and enhancements:** For each release, closing is a springboard for maintenance and enhancements; ongoing updates are critical to project success and implementation by the cancer community.

NPCR-AERRO products are reviewed, updated, enhanced, and expanded based on:

- Feedback from the cancer community
- Changes in cancer registration requirements
- Evolving technology

For example, the first release of the ePath pilot implementation identifies software and message format enhancements to be addressed in a subsequent release.

8.2.5 Monitoring and Controlling

Monitoring and Controlling takes place throughout the life of the project as plans change and products are reviewed, including schedule tracking, management of scope changes, status updates, risk and issue tracking, and review of project deliverables.

Focus for monitoring and controlling the NPCR-AERRO project:

- **Change management:** Discussing, documenting, analyzing the effects on timeline, resources, and budget, and communicating changes to internal and external stakeholders help keep project activities on track.

- **Quality management:** Products are reviewed by subject matter and technical experts within and outside the project.

8.3 Timeline

Timelines are established for each major project activity. Refer to specific activities on the NPCR-AERRO website at <http://www.cdc.gov/cancer/npcr/informatics/merp/index.htm> for updated timelines.

9 Resources

1. Williams W, Lyalin D, Wingo PA. Systems Thinking: What Business Modeling Can Do for Public Health. *Journal of Public Health Management and Practice*. 2005
2. Lyalin DS, Williams WG, Scharber WK. Applying Business Modeling Techniques to Central Cancer Registry Operations. *Journal of Registry Management* 2003; 30(2): 36-45.
3. Lyalin D, Williams W. Modeling Cancer Registration Processes with an Enhanced Activity Diagram. *Methods of Information in Medicine*
4. (Medical Problem and Document Model for Natural Language Understanding; Stephane Meystre, MD, MS and Peter J. Haug, MD; Department of Medical Informatics, University of Utah, Salt Lake City, Utah; 2003)
5. http://cdc.confex.com/cdc/nic2004/techprogram/paper_4973.htm - Immunization Registry's Immunization Information Systems; Visual Modeling of Immunization Registry Operations: Initial Efforts; Warren Williams¹, David Lyalin², Frank Caniglia³, Dennis Michaud⁴, Steven Levy⁴, Janet Kelly¹, and Ron Van Duyne¹
6. http://ncicb.nci.nih.gov/NCICB/infrastructure/cacore_overview/cadsr - caCORE and caDSR
7. caBIG – Intro to UML PDF Presentation
8. <http://www.fernuni-hagen.de/ps/pubs/SoSyM.pdf> Exploiting Practical Limitations of UML Diagrams for Model Validation and Execution; Friedrich Steimann, Heribert Vollmer

10 Appendices

Appendix 1: Domain Diagrams

A domain is an area of knowledge or activity characterized by a set of concepts and terminology understood by practitioners in the area. A domain diagram shows major business entities, their relationships and responsibilities. Unlike a data model diagram which depicts storage of information, or a workflow/process diagram which depicts the sequence of steps in a process, a domain diagram is a high-level static representation of the main “things” (entities) involved in the cancer registration process, including a description of how these “things” (entities) are related.

A domain diagram also captures a business vocabulary. It ensures that all terminology and concepts that appear in the process description are known and understood by the domain practitioners (agreed upon definitions and meaning).

How to read and interpret a domain diagram:

- Relationships between entities are visualized by connecting lines.
- Names associated with these lines describe the type of the relationship between entities.

Example: A relationship between *Hospital Cancer Registry* and *Central Cancer Registry* is shown as a connecting line with the name “reports data to”. Such a relationship should be read as “*Hospital Cancer Registry* reports data to *Central Cancer Registry*”.

Figure 11: NPCR-AERRO Hospital: Domain Diagram

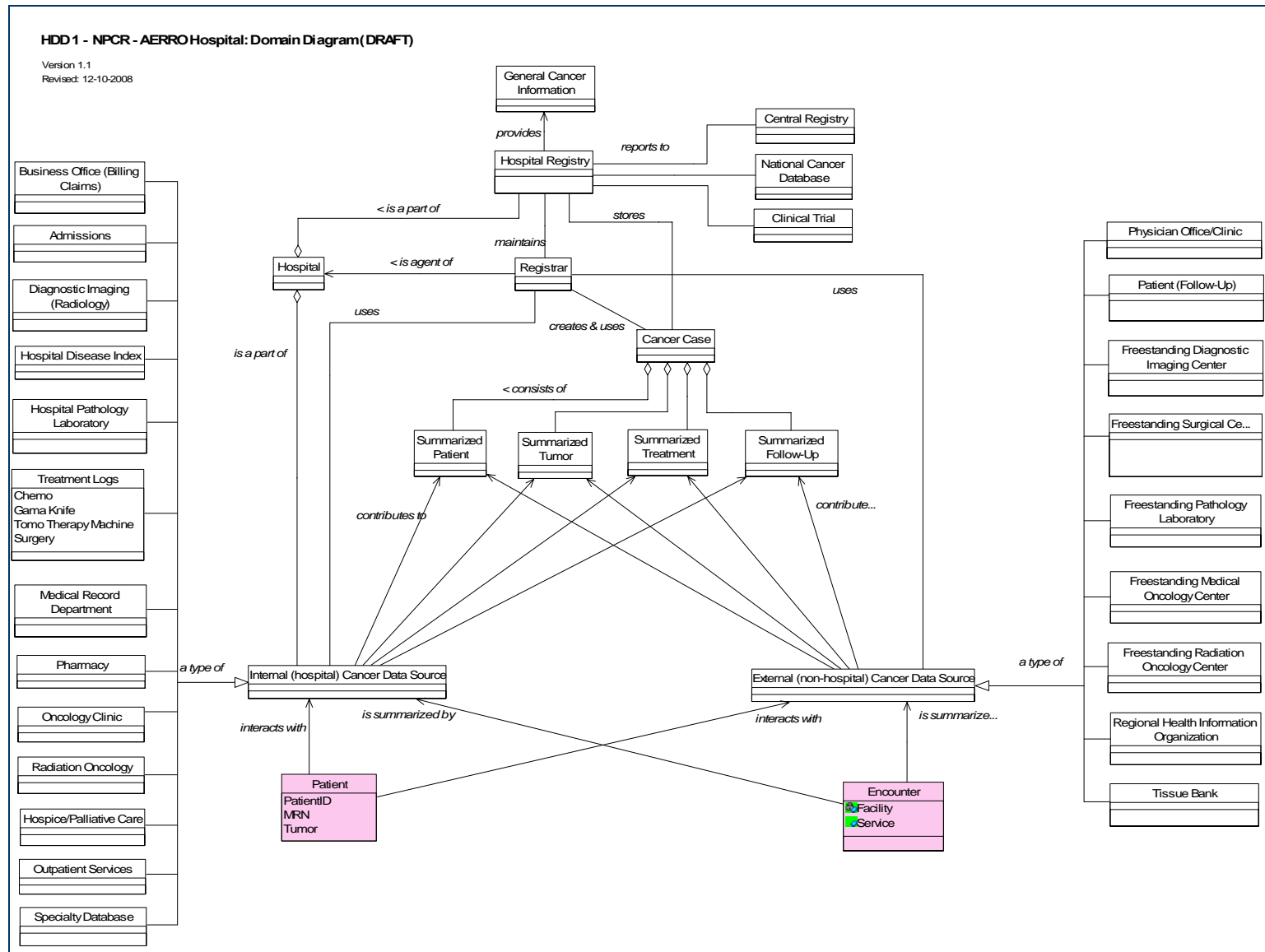
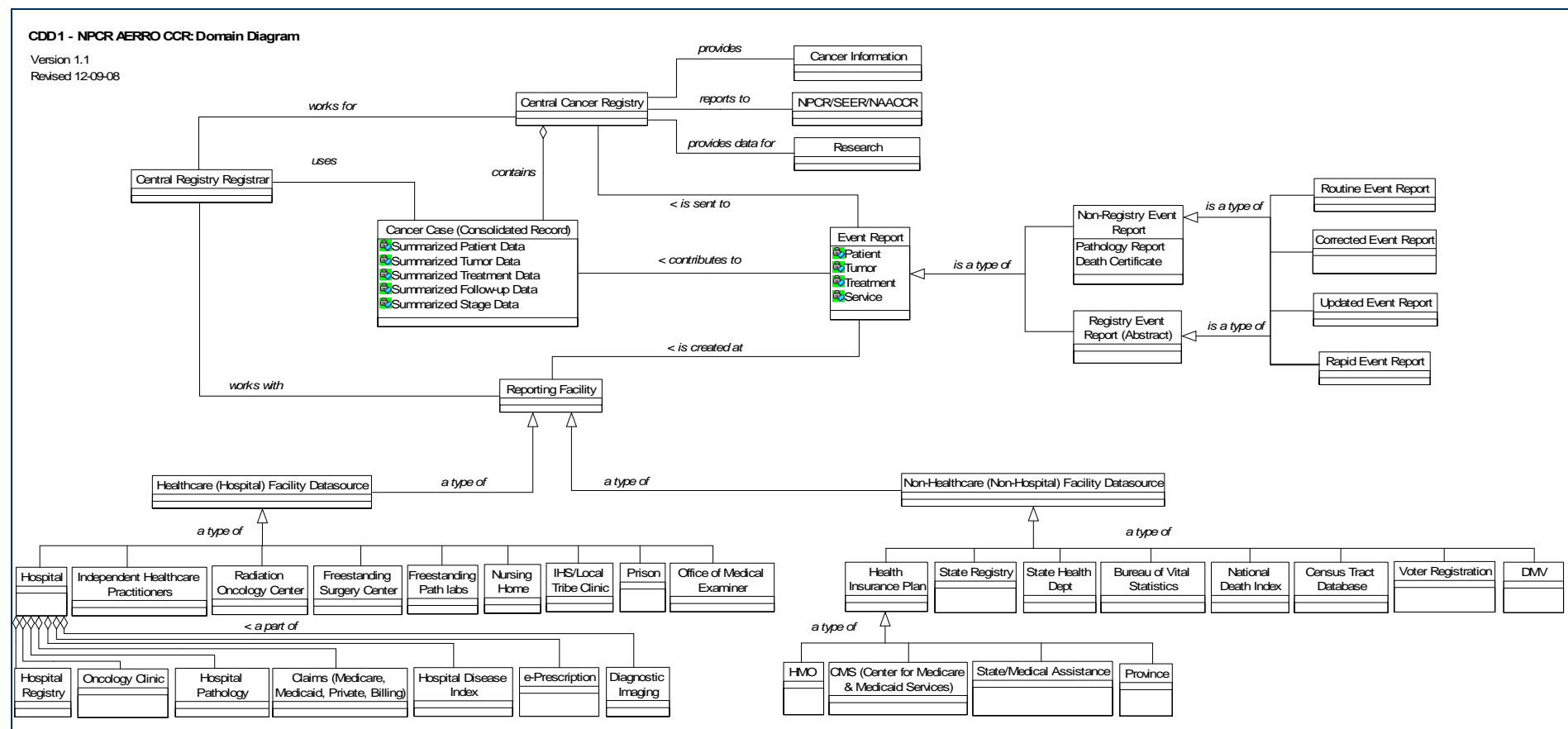


Figure 12: NPCR-AERRO CCR: Domain Diagram



Appendix 2: Use Case Diagrams

A use case diagram:

- Presents a high-level view of how the system is used as viewed from an outsider's (actor's) perspective.
- Visually depicts the behavior of the system.
- May depict all or some of the use cases of a system.
- Can be used during analysis to capture system requirements and understand how the system should work.

Operations Use Case Diagram:

The NPCR-AERRO Operations Use Case Diagram shows the grouping of cancer registry (Hospital or Central) operations under packages. A package is a container, represented as a folder, which can contain model elements such as use case and other packages.

Business Use Case Diagram:

The NPCR-AERRO Business Use Case Diagram shows the business process of a cancer registry (Hospital or Central) and its interaction with business workers and business actors. A business worker is one who acts within the system, performs the processes and interacts with other business workers and business actors. A business actor is one who plays a role in relation to the business in the business environment, affecting it externally. In the diagram below, the actors are performing the different functions of Cancer Registry (Hospital or Central). The outcome of the actions performed by actors on the functions is utilized by the recipients.

This section includes:

1. Hospital Business Use Case Diagram
2. Hospital Operations Use Case Diagram
3. CCR Business Use Case Diagram
4. CCR Operations Use Case Diagram

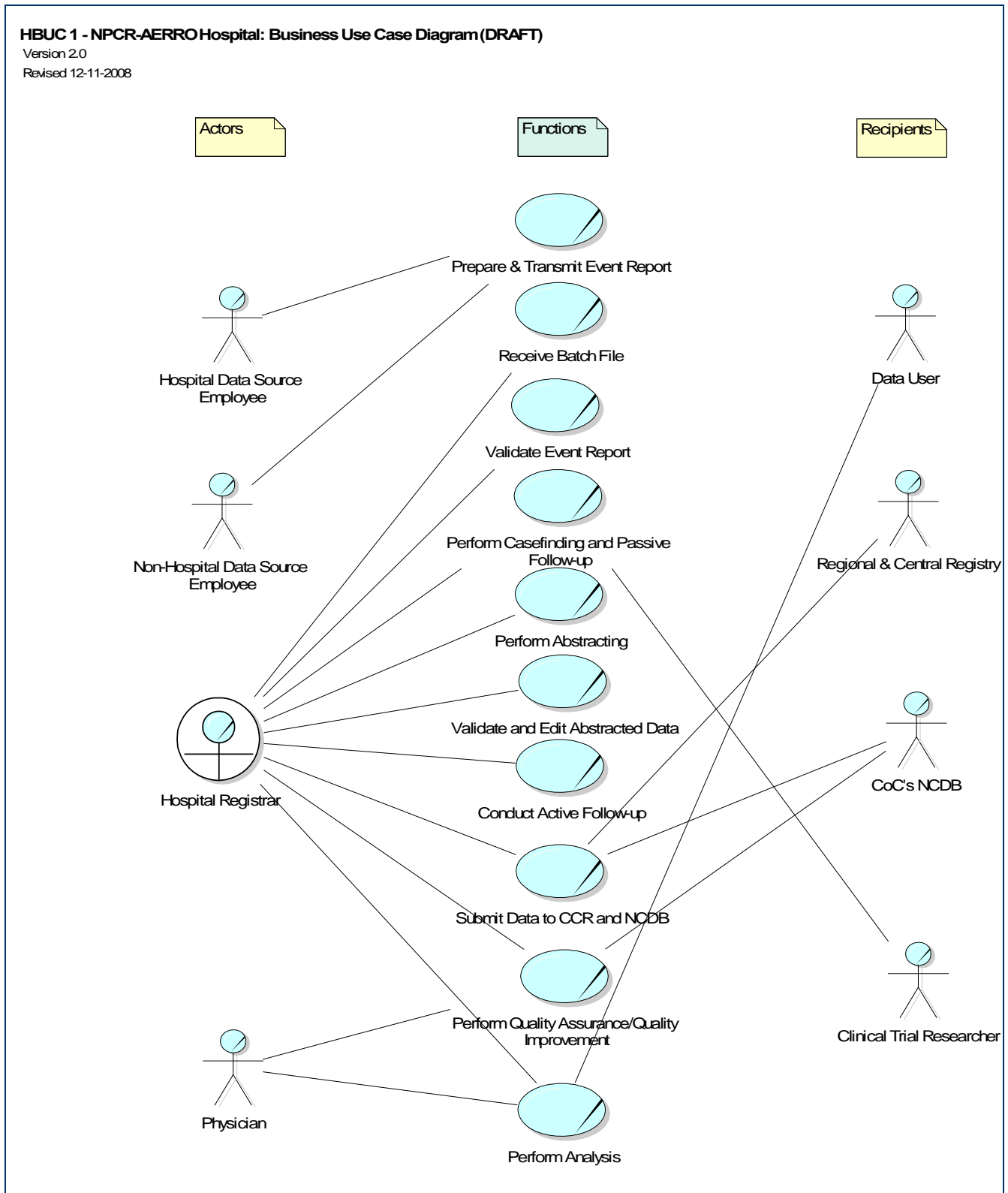
Figure 13: NPCR-AERRO Hospital: Business Use Case Diagram

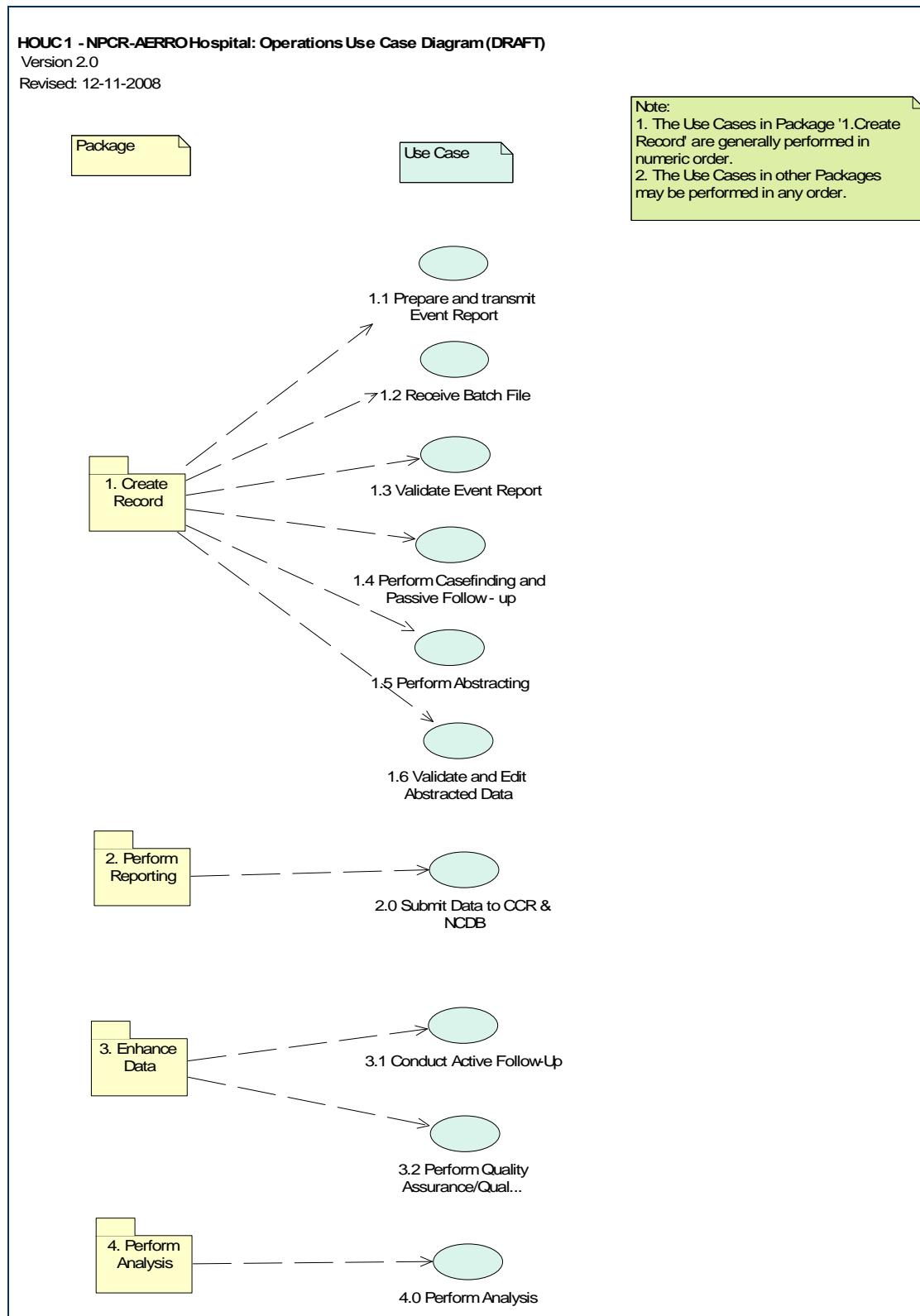
Figure 14: NPCR-AERRO Hospital: Operations Use Case Diagram

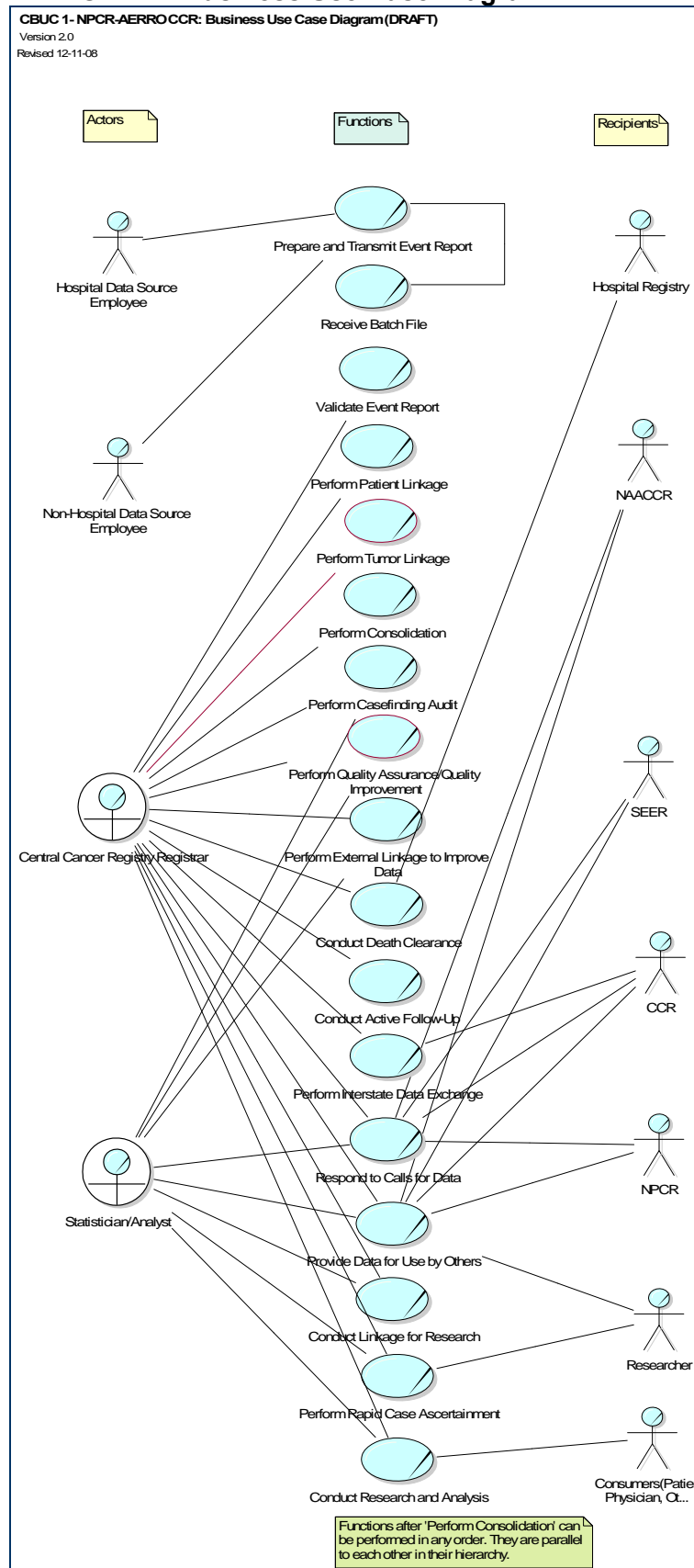
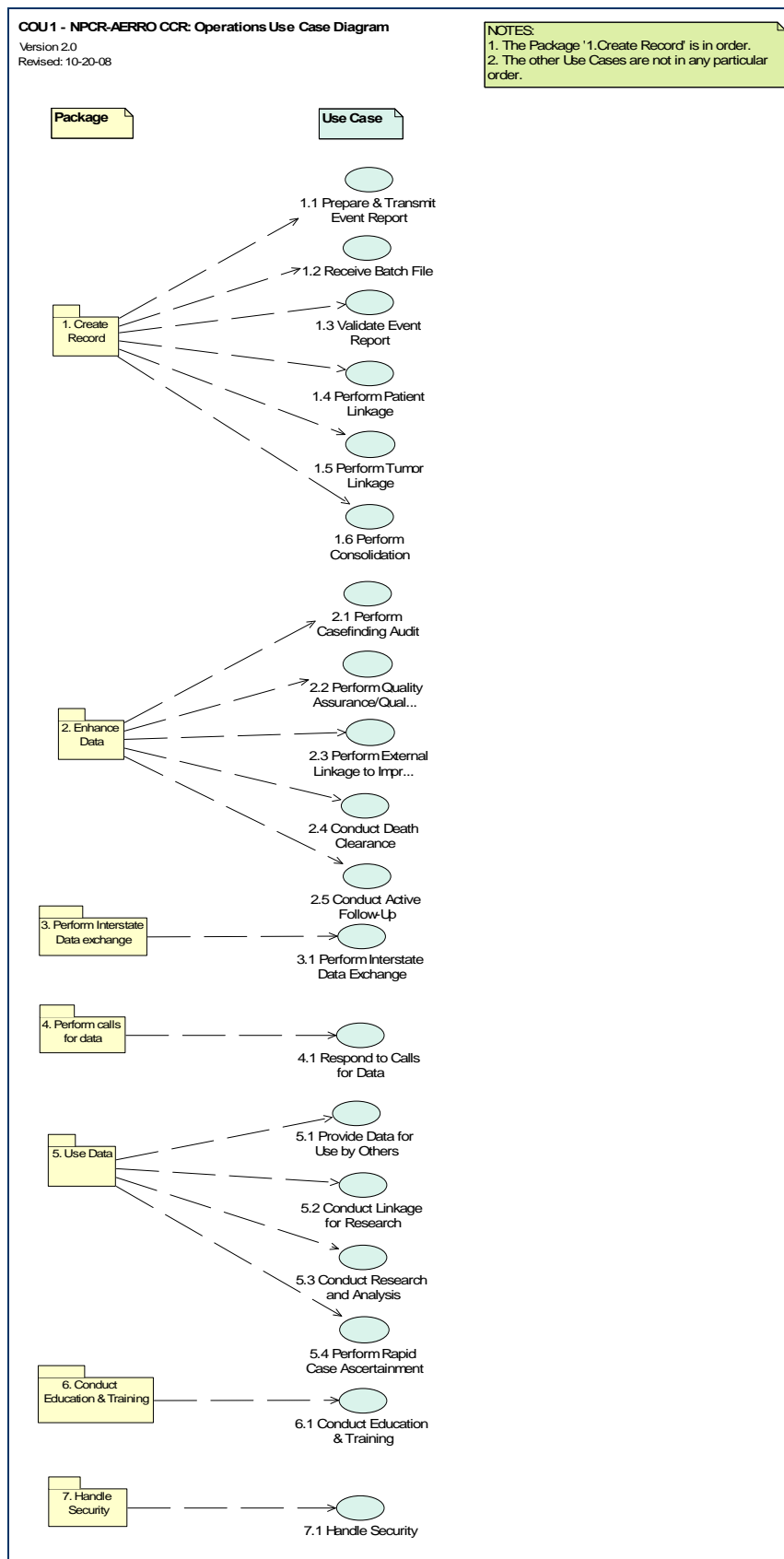
Figure 15: NPCR-AERRO CCR: Business Use Case Diagram

Figure 16: NPCR-AERRO CCR: Operations Use Case Diagram

11 References

- ¹ See the NPCR website for details on the purpose and rationale for cancer registries. Website available at: www.cdc.gov/cancer/npcr.
- ² Public Law 102-515. *Cancer Registries Amendment Act*. Available at <http://www.cdc.gov/cancer/npcr/npcrpdfs/publaw.pdf>.
- ³ Public Law 107-260. *Benign Brain Tumor Cancer Registries Amendment Act*. Available at http://www.cdc.gov/cancer/npcr/training/pdfs/Amendment_Act.pdf.
- ⁴ National Health Information Network (NHIN). Available at <http://www.hhs.gov/healthit/healthnetwork/background/>.
- ⁵ Centers for Disease Control and Prevention (CDC). *Cancer Surveillance System Rationale and Approach*. CDC, National Program of Cancer Registries (NPCR) website. Available at <http://www.cdc.gov/cancer/npcr/training/css.htm>.
- ⁶ Refer to each state's Public Laws.
- ⁷ United States Cancer Statistics. CDC Website. Available at <http://apps.ncccd.cdc.gov/uscs/>.
- ⁸ Cancer Control Planet website: <http://www.cancercontrolplanet.org/>.*
- ⁹ Cancer Statistics Review. National Cancer Institute Website. Available at <http://www.seer.cancer.gov/csr/>.
- ¹⁰ Cancer in North America. NAACCR Website. Available at http://www.naacr.org/index.asp?Col_SectionKey=11&Col_ContentID=50.
- ¹¹ International Union Against Cancer (UICC). *World Cancer Declaration, July 2006*. UICC website. Available at http://www.uicc.org/index.php?option=com_content&task=view&id=16528&Itemid=569.*
- ¹² International Agency for Research on Cancer (IARC). IARC website. Available at <http://www.iarc.fr/>.*
- ¹³ International Association of Cancer Registries. IACR website. Available at <http://www.iacr.com.fr/>.*
- ¹⁴ International Health Terminology Standards Development Organisation (IHTSDO). IHTSDO website. Available at <http://www.ihtsdo.org>.*
- ¹⁵ The White House. *Promoting Innovation and Competitiveness: President Bush's Technology Agenda*. The White House, President George. W. Bush website. Available at http://www.whitehouse.gov/infocus/technology/economic_policy200404/chap3.html.
- ¹⁶ U.S. Department of Health and Human Services (HHS). *Health Information Technology*. HHS website. Available at <http://www.hhs.gov/healthit/healthnetwork/background/>.
- ¹⁷ Cancer Surveillance System Rationale and Approach. Available at <http://www.cdc.gov/cancer/npcr/training/css.htm>.
- ¹⁸ Wingo P, Howe H, Thun M, Ballard-Barbash R, et.al. A national framework for cancer surveillance in the United States . *Cancer Causes and Control*; 16:151-170, 2005.
- ¹⁹ American College of Surgeons Commission on Cancer. *Cancer Program Standards*, 2004. Available at <http://www.facs.org/cancer/coc/cocprogramstandards.pdf>.*
- ²⁰ Griffen A, Rogers C. 2007 April. Implementing Automated Pathology Report Case Identification in a Cancer Registry. Presented at the NCRA Annual Conference; Las Vegas, Nevada.
- ²¹ Hutchison C., et al, 1997. *Cancer Registry Management: Principles and Practice*. 2nd Edition. National Cancer Registrars Association, Inc.
- ²² Hall, H. Irene, PhD, et al, "Reporting cutaneous melanoma to cancer registries in the United States Reporting", *Journal of the American Academy of Dermatology*, Volume 49, Number 4.

-
- ²³ Cole N. MCR and WebPlus: Melanoma Reporting. Presentation at the NAACCR 2006 Annual Conference. NAACCR Website. Available at: http://www.naacccr.org/filesystem/pdf/2006AC_Cole.pdf.
- ²⁴ Public Health Data Standard Consortium. White Paper: Building a Roadmap for Health Information systems Interoperability for Public Health. IHE Website. Available at www.ihe.net/Technical_Framework/upload/IHE-PHDSC_Public_Health_White_Paper_2008-07-29.pdf.
- ²⁵ Ibid.
- ²⁶ U.S. Department of Health and Human Services (HHS). *Health Information Technology*. Final Report: *Defining Key Health Information Technology Terms*. HHS website. Available at http://www.os.dhhs.gov/healthit/documents/m20080603/10_2_hit_terms.pdf
- ²⁷ Ibid, p. 17.
- ²⁸ Ibid, p. 19.
- ²⁹ Ibid, p. 16.
- ³⁰ Sittig DF; Stud Health Technol Inform. 2004;107(Pt 2):1379-82: Potential impact of advanced clinical information technology on healthcare in 2015.
- ³¹ Shortliffe Edward H.; Sondik Edward J.; CCC, Cancer causes & control ; 2006, vol. 17, no7, pp. 861-869: The public health informatics infrastructure: anticipating its role in cancer.
- ³² North American Association of Central Cancer Registries (NAACCR). NAACCR website. Available at www.naacccr.org/filesystem/pdf/E-Path%20Reporting%20Guidelines_FINAL_01-29-07%20.pdf.*
- ³³ Health Level Seven (HL7). HL7 website. Available at www.hl7.org.*
- ³⁴ Regenstrief Institute, Inc. website. Available at <http://www.regenstrief.org/medinformatics/loinc/>.*
- ³⁵ U.S. National Library of Medicine – National Institutes of Health (NLM) NLM website. Available at http://www.nlm.nih.gov/research/umls/Snomed/snomed_main.html.*
- ³⁶ DesRoche CM, et al. Electronic Health Records in Ambulatory Care – A National Survey of Physicians. The New England Journal of Medicine. 359;1 July 3, 2008. www.nejm.org
- ³⁷ American Association of Family Physicians. AAFP Survey Shows Number of FPs Using EHRs Continues to Climb [7/13/2007] AAFP Website. Available at <http://www.aafp.org/online/en/home/publications/news/news-now/practice-management/20070713ehrsurvey.html>
- ³⁸ Ashish K. Jha, et al. Use of Electronic Health Records in U.S. Hospitals. Published at www.nejm.org March 25, 2009 (10.1056/NEJMsa0900592)
- ³⁹ Business Improvement Architects (BIA). BIA website. Available at <http://www.bia.ca/articles/chng-managing-resistance.htm>.*
- ⁴⁰ U.S. Department of Health and Human Services (HHS). *Health Information Technology*. HHS website. Available at <http://www.hhs.gov/healthit/ahic.html>
- ⁴¹ American National Standards Institute (ANSI). ANSI website. Available at http://www.ansi.org/standards_activities/standards_boards_panels/hisb/hitsp.aspx?menuid=3.*
- ⁴² Modeling Electronic Reporting Project (MERP) Summary Technical Report: Information Processing Within The Virginia Commonwealth University Massey Cancer Center Cancer Registry System. CDC Website. Available at <http://ftp.cdc.gov/pub/Software/RegistryPlus/MERP/VCU%20Summary%20Technical%20Report.pdf>.
- ⁴³ Requirements Findings NPCR-MERP Hospital Strategic Assessment and Modeling Session (SAMS). Centers for Disease Control and Prevention Website. Available at http://www.cdc.gov/cancer/npcr/npcrpdfs/MERP_sams_hospital_2006_02.pdf.

-
- ⁴⁴ Public Health Data Standards Consortium. "Building a Roadmap for Health Information Systems Interoperability for Public Health". 2008.
http://www.ihe.net/Technical_Framework/upload/IHE_PHDSC_Public_Health_White_Paper_2007_10_11.pdf.*
- ⁴⁵ Cockburn M, Swetter S, Peng D, Keegan T, Deapen D, Clark C. Melanoma underreporting: Why does it happen, how big is the problem and how do we fix it? *Journal of American Academy of Dermatology*. Volume 9, Number 6. pg 1081-1085.
- ⁴⁶ Hall IE, et al. The Reporting of Cutaneous Melanoma to Cancer Registries in the United States. NAACCR Website. Available at: <http://www.naacr.org/filesystem/pdf/melanomadraft8.PDF>.
- ⁴⁷ Lyalin DS, Williams WG, Scharber WK. Applying Business Modeling Techniques to Central Cancer Registry Operations. *Journal of Registry Management* 2003;30(2):36-45.
- ⁴⁸ U.S. National Institutes of Health – National Cancer Institute Center for Bioinformatics (NCICB). NCICB website. Available at http://ncicb.nci.nih.gov/NCICB/infrastructure/cacore_overview
- ⁴⁹ Extreme Programming website. Available at <http://www.extremeprogramming.org/>.*
- ⁵⁰ U.S. National Institutes of Health – National Cancer Institute Center for Bioinformatics (NCICB). NCICB website. Available at http://ncicb.nci.nih.gov/infrastructure/cacore_overview/caBIO/software_process
- ⁵¹ U.S. National Institutes of Health – National Cancer Institute Surveillance Epidemiology and End Results (SEER). SEER website. Available at <http://seer.cancer.gov/seerdms/manual/chap01.overview.final.pdf>
- ⁵² U.S. National Institutes of Health – National Cancer Institute Surveillance Epidemiology and End Results (SEER). SEER website. Available at <http://seer.cancer.gov/seerdms/path/approach.html>
- ⁵³ North American Association of Central Cancer Registries (NAACCR). NAACCR website. Available at http://www.naacr.org/filesystem/pdf/E-Path%20Reporting%20Guidelines_FINAL_01-29-07%20.pdf.*
- ⁵⁴ Department of Computer Science at Louisiana State University website. Available at <http://bit.csc.lsu.edu/~chen/chen.html>.*
- ⁵⁵ Wikipedia: The Free Encyclopedia website. *Entity-relationship model*. Wikipedia website. Available at http://en.wikipedia.org/wiki/Entity-relationship_model.*
- ⁵⁶ Agile Modeling. *Effective Practices for Modeling and Documentation*. Agile Modeling website. Available at <http://www.agilemodeling.com/>.*
- ⁵⁷ International Business Machines (IBM) Rational Unified Process. IBM website. Available at http://www.ibm.com/developerworks/rational/library/content/03July/1000/1251/1251_bestpractices_TPO26B.pdf.*
- ⁵⁸ Object Management Group (OMG). OMG website. Available at http://www.omg.org/gettingstarted/what_is_uml.htm.*
- ⁵⁹ Centers for Disease Control and Prevention (CDC). CDC website. Available at <http://www2.cdc.gov/cdcup>
- ⁶⁰ UML in a Nutshell: A Desktop Quick Reference (Nutshell Handbook)
- ⁶¹ U.S. National Institutes of Health – National Cancer Institute Cancer Biomedical Informatics Grid (CaBIG). CaBIG website. Available at https://cabig.nci.nih.gov/workspaces/VCDE/Documents/Useful_Presentations/Model%20Driven%20Architecture/caBIO_UML.pdf
- ⁶² FernUniversität in Hagen. *Exploiting Practical Limitations of UML Diagrams for Model Validation and Execution*; Friedrich Steimann, Heribert Vollmer. FernUniversität in Hagen website. Available at <http://www.fernuni-hagen.de/ps/pubs/SoSyM.pdf>.*

-
- ⁶³ The UML Reference Manual by Rumbaugh, Jacobson & Booch (SECOND EDITION) - published by Addison-Wesley
- ⁶⁴ Object Management Group (OMG). OMG website. Available at http://www.omg.org/gettingstarted/what_is_uml.htm.*
- ⁶⁵ College of American Pathologists (CAP). CAP website. Available at http://www.cap.org/apps/cap.portal?_nfpb=true&cntvwrPtl_t_actionOverride=%2Fportlets%2FcontentViewer%2Fshow&_windowLabel=cntvwrPtl_t&cntvwrPtl_t%7BactionForm.contentReference%7D=snomed%2Fabout_ecc.html&_state=maximized&_pageLabel=cntvwr.*
- ⁶⁶ North American Association of Central Cancer Registries (NAACCR). NAACCR website. Available at http://www.naaccr.org/index.asp?Col_SectionKey=7&Col_ContentID=501.*
- ⁶⁷ North American Association of Central Cancer Registries (NAACCR). NAACCR website. Available at http://www.naaccr.org/index.asp?Col_SectionKey=7&Col_ContentID=501.*
- ⁶⁸ Centers for Disease Control and Prevention (CDC). *CDC Unified Process - A Common Project Delivery Framework v3.0*, CDC website. Available at <http://www2.cdc.gov/cdcup/>
- ⁶⁹ Centers for Disease Control and Prevention (CDC). *CDC UP Project Framework*, CDC website. Available at <http://www2.cdc.gov/cdcup/library/pmg/default.htm>